# Privacy Year in Review: Privacy Impact Assessments, Airline Passenger Pre-Screening, and Government Data Mining

SAYAKA KAWAKAMI & SARAH C. MCCARTY*

ABSTRACT

*This article discusses the privacy issues associated with personal information collected and used by the government. The E-Government Act of 2002 requires government agencies to use Privacy Impact Assessments to address the privacy concerns of new government information collection schemes, for instance, with government websites and airline passenger pre-screening systems. However, even with the privacy assessments, airline passenger pre-screening systems have encountered some questions as to whether adequate measures exist to protect passenger privacy. This article will discuss the developments in 2004 associated with passenger pre-screening systems. Privacy issues are also implicated by the government's use of data mining. Data mining allows the government to analyze data collected, including personal information, and convert it into a useful form. This article will discuss the uses of data mining, such as for terrorism prevention, on the federal and state government levels as well as the privacy concerns associated with government data mining.*

## INTRODUCTION

The Fourth Amendment provides that citizens shall be free from unreasonable government intrusion. This provision has been interpreted by the courts as a right to privacy against government intrusion.[1] The collection of information by the government and the use of that information are seen by some as an intrusion by the government, thus requiring privacy protection. The major privacy law that governs government information collection is the Privacy Act of 1974[2], which regulates the federal government's use of personal information by restricting and monitoring agencies' collection, disclosure, and use of personal information. The intersection of

---

* The authors are both J.D. candidates at The Ohio State University Moritz College of Law, class of 2006. Sayaka Kawakami holds a B.A. in political science and communications from the University of Washington. Sarah McCarty holds both a bachelor's and master's degree in business administration from Ohio University.

[1] U.S. CONST. amend. IV, http://www.usconstitution.net/const.html#Am4.

[2] Privacy Act of 1974, 5 U.S.C. § 552.

government information collection and privacy encompasses a broad variety of topics and programs. This article focuses on the most pertinent topics affecting privacy today and the new developments in the field over the past year. Privacy impacts with regard to government websites, Transportation Security Administration data collection and usage, and restrictions on government data mining have been recent sources of change.

Both new electronic technology and the threat of terrorism in the aftermath of the terrorist attacks on September 11, 2001, have led the government to adopt new information collection schemes, which create new privacy concerns among citizens. In order to address such privacy concerns, Congress adopted the E-Government Act of 2002[3], requiring government agencies to submit privacy impact assessments (PIAs), which identify privacy concerns and ensure privacy protection. The Act also mandates that the Office of Management and Budget issue guidelines to federal agencies for their websites to ensure that privacy of citizens who access their sites is protected. Even with these new government website privacy policies and PIA procedures, some of the concerns over individual privacy still remain unsolved. This problem can be seen in the implementation of new programs, such as CAPPS II and Secure Flight, both of which were proposed to identify terrorists using computer-assisted passenger screening systems.

Government data mining is present at both the federal and state levels. The Total Information Awareness Program (TIA) emerged after September 11, 2001, to create technologies for terrorism prevention but was quickly disassembled. Other federal data mining systems were spawned from the TIA or originally created, and are currently used by fifty two federal agencies. The passage of the Intelligence Reform and Terrorism Prevention Act of 2004[4], has granted greater authority to pursue government data mining, but also placed more severe checks on the program. The MATRIX program has provided data mining to the states. Currently, only five states are participating in this federally funded, privately operated system. As the technologies used in data mining increase, the conflicts with privacy protection increase. Two primary issues that have arisen are the use of private database contractors to supply information to the government, and how to balance the usefulness and efficiency of data mining against the need for privacy protection.

---

[3] E-Government Act of 2002, Pub. L. No. 107-347 (codified as amended at 44 U.S.C. § 36).

[4] Intelligence Reform and Terrorism Prevention Act, Pub. L. No. 108-458 (codified at 50 U.S.C. § 401).

This article focuses on the details of these debates and challenges. Each topic is analyzed to determine the new developments in the federal and state governments, recent cases and controversies in the subject area, interesting topics or debates, and a brief analysis of the future direction of government information collection. The first part focuses on Privacy Impact Assessments and Computer-Assisted Passenger Screening Programs. It focuses its discussion on the role of Privacy Impact Assessments on government websites and CAPPS II. The article then turns to the developments in these areas during 2004, such as the Crew Vetting Program, No-Fly Lists, and Secure Flight. The second part of the article covers government data mining. This section discusses data mining efforts and privacy concerns on the federal level regarding the Total Information Awareness Program and the Intelligence Reform and Terrorism Prevention Act of 2004 and on a state level by looking at the MATRIX program. The article provides analysis of how the government can balance privacy with useful data mining, and how privacy is affected by data mining from private institutions. The second part concludes with a projection into the future of government data mining.

## PRIVACY IMPACT ASSESSMENTS AND COMPUTER-ASSISTED PASSENGER SCREENING PROGRAMS

In the last few years, the federal government has established the use of Privacy Impact Assessments (PIAs), which were created to answer privacy concerns of new government information collection schemes. This section discusses PIAs mandated under the E-Government Act of 2002. Government agencies have not received much criticism concerning the privacy policies of government websites even though more and more people use the Internet as a tool to contact the agencies. On the other hand, new measures proposed by the Transportation Security Administration (TSA) on computer-assisted airline passenger pre-screening systems, such as CAPPS II and Secure Flight, have spurred some criticism from the public despite the fact that the TSA issued PIAs to ensure that all privacy issues are addressed in implementing those measures. First, this section will discuss the E-government Act of 2002 and the authority given to the Office of Management and Budget (OMB) to implement the privacy provision of the Act. Then, this article will discuss the application of the Act to government websites and computer-assisted airline passenger pre-screening systems.

## I. PRIVACY IMPACT ASSESSMENT (PIA)

Federal agencies are required by law to issue a PIA whenever they collect information from the public by electronic means. The following subsection gives a brief introduction to PIAs.

### A. BACKGROUND

In 2003, forty million Americans visited a government website for information, up fifty percent from the previous year.[5] The increased use of electronic tools by citizens led Congress to pass the E-Government Act of 2002,[6] which went into effect on April 17, 2003. The purpose of the Act is to foster the use of information technology (IT) for the agency's business among citizens by making it more citizen-oriented and user-friendly.[7] In order to achieve this purpose, Congress believed that it was important to protect the privacy of citizens when they interact with the federal government through the use of IT.[8]

Section 208 of the Act mandates that the OMB issue guidance on implementing the privacy provision of the Act.[9] The section specifically requires that each government agency "conduct privacy impact assessments before developing or producing information technology that collects, maintains, or disseminates information that is in an identifiable form."[10] This provision also applies in a case where

---

[5] Pamela M. Prah, *E-Government Use Up 50 Percent, Survey finds*, STATELINE.ORG (May 25, 2004), *at* http://www.stateline.org/live/ViewPage.action?siteNodeId=136&languageId=1&contentId=15659.

[6] E-Government Act of 2002, Pub. L. No. 107-347 (codified as amended at 44 U.S.C. § 36), *available at* http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:pub1347.107.pdf.

[7] Memorandum from Joshua B. Bolten, Director, Office of Management and Budget, Implementation Guidance for the E-Government Act of 2002, M-03-18 (Aug. 1, 2003), *available at* http://www.whitehouse.gov/omb/memoranda/m03-18.pdf.

[8] Memorandum from Joshua B. Bolten, Director, Office of Management and Budget, OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002, M-03-22 (Sept. 26, 2003), *available at* http://www.whitehouse.gov/omb/memoranda/print/m03-22.html.

[9] *Id.*

[10] E-Government Act of 2002, *supra* note 3.

an agency initiates a new collection of information that "will be collected, maintained, or disseminated using information technology or includes any information in an identifiable form permitting the physical or online contacting of a specific individual."[11] In addition, agencies are required to update PIAs when a system change creates new privacy risks.[12] Those changes include conversion from paper-based records to electronic systems, changes in collecting information from anonymous to non-anonymous, significant system management changes, significant merging, centralization, and matching with other databases, matching of information with commercially available information, alteration in the character of the data, new public access with passwords and digital certificates, new interagency uses, and changes in internal flow or collection of information.[13]

The OMB guidance defines a PIA as "an analysis of how information is handled by federal agencies."[14] The purposes of PIAs are to ensure information handling conforms to applicable legal, regulatory, and policy requirements on privacy, to determine the risks and effects on privacy of collecting, maintaining, and disseminating information of citizens in identifiable form in an electronic information system, and to examine and evaluate protections and alternative processes for handling information to mitigate potential privacy risks.[15] The OMB specifically requires that a PIA address the following: what information is to be collected, why the information is being collected, the intended use by the agency of the information, with whom the information will be shared, how the information will be secured, whether a system of records required under the Privacy Act of 1974 is being crafted, and whether a privacy policy is in machine-readable format.[16]

---

[11] Bolten, *supra* note 7.

[12] *Id.*

[13] *Id.*

[14] *Id.*

[15] *Id.*

[16] Bolten, *supra* note 7.

### i. GOVERNMENT WEBSITES

Since 1999, the OMB has been requiring federal agencies to state the nature, purpose, use, and sharing of information collected on their websites in their "Privacy Policy."[17] Particularly, on "cookies," federal agencies could only use them or other automatic means of collecting personal information if they give a notice of those activities.[18] The 2003 OMB Memorandum, mandated under the E-Government Act of 2002, additionally requires that agencies implement changes regarding privacy to their websites by December 15, 2003.[19] Under the new mandate, agencies must place notices on their websites explaining agency information handling practices labeled as "Privacy Policy."[20] The Privacy Policy must address: (1) consent to collection and sharing, and (2) rights under the Privacy Act of 1974 or other privacy laws.[21] Furthermore, agencies must adopt machine-readable technology so that site users are alerted automatically when site privacy practices do not match their personal privacy preferences.[22] Additionally, agencies are required to develop a timetable for translating their privacy policy into a standardized machine-readable format.[23] However, these new policies do not apply to information other than that on government agency intranet sites that are not accessible by the public and national security systems.[24]

---

[17] Memorandum from Jacob J. Lew, Director, Office of Management and Budget, Privacy Policies on Federal Web Sites, M-99-18 (Jun. 2, 1999), *available at* http://www.whitehouse.gov/omb/memoranda/m99-18.html.

[18] Memorandum from Jacob J. Lew, Director, Office of Management and Budget, Privacy Policies and Data Collection on Federal Web Sites, M-00-13 (Jun. 22, 2000), *available at* http://www.whitehouse.gov/omb/memoranda/m00-13.html. A cookie is a file on a Web user's hard drive that is used by Web sites to record data about the user.

[19] Bolten, *supra* note 7.

[20] *Id.*

[21] Bolten, *supra* note 7. Such statutes include the Health Insurance Portability and Accountability Act of 1996, and the I.R.S. Restructuring and Reform Act of 1998.

[22] Bolten, *supra* note 7.

[23] *Id.*

[24] *Id.*

### ii. STATE GOVERNMENT WEBSITES

Currently, thirty-one states place some sort of privacy policy notice on their websites.[25] At least sixteen of those states have a statute requiring government websites to establish privacy policies and procedures, or to incorporate machine-readable privacy policies into their websites. These states are: Arizona, Arkansas, California, Colorado, Delaware, Iowa, Illinois, Maine, Maryland, Michigan, Minnesota, Montana, New York, South Carolina, Texas, and Virginia.[26]

### iii. COMPUTER-ASSISTED PASSENGER SCREENING SYSTEM

#### a. BACKGROUND

Computer-assisted airline passenger pre-screening systems involve the collection of passenger information using electronic databases. Thus, issuance of a PIA is mandated to the Transportation Security Administration (TSA) by the E-Government Act of 2002 in order to ensure that a proper privacy policy protects the TSA from misusing such information. The privacy policy of each computer-assisted airline passenger pre-screening system is discussed in the following.

#### b. CAPPS

In fear of the increasing threat of terrorism, the federal government had been using a computer-based passenger screening system called the Computer-Assisted Passenger Pre-screening System (CAPPS) since 1998. CAPPS was operated by U.S. commercial airlines, each of which kept separate computer systems. CAPPS analyzed information in passenger name records (PNRs).[27] The information was collected

---

[25] NAT'L CONFERENCE OF STATE LEGISLATURES, LEGISLATIVE WEB SITE PRIVACY POLICIES (2005), *at* http://www.ncsl.org/programs/lis/nalit/legprivacypol.htm (last updated Jan. 4, 2005).

[26] NAT'L CONFERENCE OF STATE LEGISLATURES, STATE LAW RELATED TO INTERNET PRIVACY (2005), *at* http://www.ncsl.org/programs/lis/privacy/eprivacylaws.htm (last updated Jan. 3, 2005).

[27] Information collected for PNRs varied among airlines. It may include passenger name, reservation date, travel agency or agent, travel itinerary information, form of payment, flight number, and seating location.

when passengers made flight reservations with the airlines.[28] Under CAPPS, passengers who fit a certain behavioral profile, such as those who buy a one-way ticket or pay by cash, were subjected to extra screening.[29] Under the CAPPS system, half of the hijackers involved in the terrorist attacks on September 11th were flagged, but it did not prevent the attacks.[30] The TSA later pointed out that the amount of information in PNRs under CAPPS was very limited because of the unclassified nature of the system.[31]

### c. CAPPS II

In November 2001, Congress passed the Aviation and Transportation Security Act,[32] which mandated that the newly-created TSA within the Department of Transportation[33] replace CAPPS with a new computer-assisted passenger pre-screening system in order to evaluate all passengers before they board an aircraft.[34] The task of developing the new program was given to the Office of National Risk Assessment, which eventually developed a program known as CAPPS II.[35]

CAPPS II would take the procedure away from the airlines and hand it over to the TSA. Upon reservation of an air flight, the airline

---

[28] TRANSPORTATION SECURITY ADMINISTRATION, REPORTS, FORMS, AND RECORD KEEPING REQUIREMENTS: AGENCY INFORMATION COLLECTION ACTIVITY UNDER OMB REVIEW; SECURE FLIGHT TEST PHASE, Docket No. TSA-2004-19160, 2-3, *available at* http://www.epic.org/privacy/airtravel/sf_pra_9.21.04.pdf.

[29] Sara Kehaulani Goo and Robert O'Harrow, Jr., *TSA Readies Revised Aviation Screening*, THE WASH. POST, Aug. 26, 2004, at A12, *available at* http://www.washingtonpost.com/ac2/wp-dyn/A33830-2004Aug25?language=printer.

[30] *Id.*

[31] TRANSPORTATION SECURITY ADMINISTRATION, *supra* note 28, at 3.

[32] Pub.L. No. 107-71 (codified as amended at 49 U.S.C. § 114), *available at* http://usinfo.state.gov/usa/infousa/laws/majorlaw/107-71.pdf.

[33] Under the Homeland Security Act of 2002, the TSA became a component of the Department of Homeland Security on March 1, 2003. 6 U.S.C. § 203(5) (2005).

[34] U.S. GENERAL ACCOUNTING OFFICE, AVIATION SECURITY: COMPUTER-ASSISTED PASSENGER PRESCREENING SYSTEM FACES SIGNIFICANT IMPLEMENTATION CHALLENGES, GAO-04-385, 6 (Feb. 2004), *available at* http://www.gao.gov/new.items/d04385.pdf.

[35] *Id.*

requests that the passenger provide his or her name, home address, home phone number, and date of birth.[36] Then, the airline enters the information into the PNR[37] which is sent electronically to CAPPS II, managed by the TSA.[38] Sometime before the day of the flight, CAPPS II would request an identity authentication from commercial data providers.[39] Based on an identity authentication score the commercial database identified, CAPPS II would conduct risk assessments using government databases, including classified and intelligence data, to categorize the passenger into one of three classes: an acceptable risk, unknown risk, or unacceptable risk.[40] When the passenger checks in for a flight at the airport, the passenger's risk class will be transmitted from the CAPPS II to the check-in counter.[41] Passengers who are an acceptable or unknown risk will receive a boarding pass encoded with their risk level so that checkpoint screeners will know the level of scrutiny required.[42] Those people with unacceptable risk would not be allowed to fly, would undergo police questioning, and risk possible arrest.[43] Those with unknown risk would be subject to extra screening by TSA agents at the airport screening gate.[44]

## B. DEVELOPMENT IN THE YEAR 2004

In 2004, the TSA's proposal of CAPPS II received criticism from both Congress and the public. This eventually led the TSA to propose and test alternative methods of computer-assisted passenger screening

---

[36] *Id.*

[37] PNR contains data related to a traveler's reservation and travel itinerary, and is contained in an air carriers reservation system.

[38] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34 at 7.

[39] *Id.*

[40] *Id.*

[41] *Id.*

[42] *Id.*

[43] ELECTRONIC FRONTIER FOUNDATION, CAPPS II: GOVERNMENT SURVEILLANCE VIA PASSENGER PROFILING, *at* http://www.eff.org/Privacy/cappsii/background.php (accessed Jan. 20, 2005).

[44] *Id.*

systems, such as the Registered Traveler Pilot Program, Crew Vetting Program, and the Secure Flight Program. The latest scheme, Secure Flight, is still unfolding.

### i. CAPPS II

CAPPS II received scrutiny from Congress in 2003. The Department of Homeland Security Appropriations Act, passed in 2003, prohibits the TSA from receiving funding for CAPPS II in the fiscal year ending September 30, 2004, until the U.S. General Accounting Office (GAO) submits a report to the Committees on Appropriations of both the Senate and House by February 15, 2004.[45] The Act mandated that the GAO address eight key issues, the first six of which show that the development and operation of CAPPS II is effectively managed and monitored, and that the system will function as intended.[46] Those six issues are: (1) establishment of an internal oversight board by the Department of Homeland Security (DHS) to monitor the development of the system; (2) assessment of database accuracy in order to avoid producing a large number of false positives; (3) creation of a test system and the demonstration of efficacy and accuracy of the system; (4) installment of operational safeguards to protect the system from abuse; (5) installment of security measures to protect the system from unauthorized access; and (6) establishment of effective oversight of the system's use and operation.[47] The last two issues involve public assurance that adequate measures exist to protect passenger privacy: (7) addressing all privacy concerns with the system; and (8) creation of a redress process for passengers to correct erroneous information.[48] On the eighth issue, the GAO must show that CAPPS II ensures that a system of due process is in place when passengers of an aircraft pose a threat and are delayed or prohibited from boarding scheduled flights.[49]

---

[45] Department of Homeland Security Appropriations Act, Pub. L. No. 108-86, *available at* http://frwebgate.access.gpo.gov/cgi-bin/useftp.cgi?IPaddress=162.140.64.21&filename=h2555rs.pdf&directory=/diskb/wais/data/108_cong_bills.

[46] *Id.*

[47] *Id.*

[48] *Id.*

[49] *Id.*

As required by the Department of Homeland Security Appropriations Act, the GAO issued an assessment of CAPPS II titled "Aviation Security: Computer-Assisted Passenger Prescreening System Faces Significant Implementation Challenges" to Congressional committees in February 2004.[50] The report identified delay in the development of CAPPS II. For example, testing and developing the initial increments of CAPPS II faced some delay partly due to the difficulty in obtaining passenger data needed to test initial increments.[51] Also, the TSA had not established a complete plan on how functionality of the system would be delivered, the schedule for delivery, and the estimated costs of CAPPS II development.[52]

Specifically, the GAO raised concerns that the TSA has not fully addressed seven of the eight issues identified by Congress under the Security Appropriations Act discussed above.[53] The first issue – establishment of an internal oversight board – was the only one fully addressed by the TSA. The TSA addressed this issue when the DHS created the Investment Review Board to review DHS's capital asset programs with contracts exceeding fifty million dollars in order to ensure that projects meet mission needs at the expected levels of cost and risk.[54]

On the second issue, there is no industry standard of accuracy assessment: each commercial provider uses different criteria to assess accuracy.[55] Also, there is no consistent system for collecting accuracy data among the government databases.[56] On the third issue, lack of data delayed demonstration of system efficacy and accuracy.[57] On the fourth and fifth issues, critical elements of the security program, including security policies, a system security plan, and certification and accreditation of the security system, had not been implemented.[58]

---

[50] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34.

[51] *Id.* at 9.

[52] *Id.* at 9-10.

[53] *Id.* at 13.

[54] *Id.* at 13-14.

[55] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34, at 14.

[56] *Id.*

[57] *Id.* at 16.

[58] *Id.* at 18.

On the sixth issue, the measures and goals for CAPPS II addressed by the TSA had not provided enough objective data necessary to conduct oversight.[59] In addition, the TSA had not fully established the system to ensure that the program would be properly monitored and evaluated.[60]

On the seventh issue on public assurance of privacy, the GAO stated that the TSA had not finalized its plan to comply with the Privacy Act of 1974 by addressing all privacy concerns.[61] The OMB guidance requires that an agency proposing to exempt a system of records from the Act must explain the reason why it wishes it to be exempted, and the TSA had not done so.[62]

According to the GAO, the TSA submitted its plans to address privacy issues in compliance with the Privacy Act.[63] For example, the TSA addressed an intention not to collect passengers' social security numbers from commercial databases and to destroy most passenger information after they have completed their travel itinerary.[64] This is consistent with the collection limitation practice of the Act.[65] Also, it proposed to prohibit commercial data providers from using information they receive from the TSA for commercial purposes.[66] This practice is consistent with the use limitation practice under the Act.[67] Furthermore, the TSA also made proposals to provide passengers with a Privacy Act notice. By doing so, the TSA sought to explain its authority for collecting their information and its principal purposes, to provide other information as the Act requires, to perform real-time auditing and testing in order to identify data quality problems, and to improve accuracy.[68]

---

[59] *Id.* at 20-21.

[60] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34, at 20-21.

[61] *Id.* at 23-24.

[62] *Id.* at 23.

[63] *Id.*

[64] *Id.*

[65] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34, at 23.

[66] *Id.* at 23-24.

[67] *Id.* at 24.

[68] *Id.*

However, the GAO noted that the TSA's plans gave rise to several privacy concerns. First, the TSA planned to exempt CAPPS II from the Privacy Act's requirements that an agency maintain only information about an individual that is relevant and necessary to accomplish a proper agency purpose.[69] The GAO stated that this policy raises concerns that the TSA may collect and maintain more information than necessary for the purpose of CAPPS II.[70] The GAO also speculated that TSA might use this information for new purposes in the future.[71] Secondly, TSA's plan to prohibit passengers from accessing their personal information obtained by CAPPS II raises the concern that inaccurate personal information would remain uncorrected in the system and continue to be used.[72]

The public raised similar concerns about CAPPS II. In Alaska, four plaintiffs filed a suit against the TSA and the DHS in the U.S. District Court of Alaska on May 24, 2004, claiming that CAPPS II would violate their constitutionally protected rights.[73] Two of four plaintiffs are U.S. citizens and Alaska residents who frequently use air travel because they live in a remote location.[74] The two other plaintiffs are Frontier Travel and Airlines Online, travel agencies that reserve commercial flights for customers in Alaska.[75] The plaintiffs claim that CAPPS II was intended to reach beyond terrorism. This was apparent by TSA's statement entitled "CAPPS II at a Glance," that passengers with outstanding warrants for a "crime of violence" would face arrest.[76] They also claimed that defendants intended to implement CAPPS II by a secret order, so that there is no way for passengers to know if their information is collected.[77] Under the

---

[69] *Id.*

[70] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 34, at 24.

[71] *Id.*

[72] *Id.*

[73] Brief for Frontier Travel at 11-12, Frontier Travel v. Transportation Security Administration (D. Alaska 2004), *available at* http://209.123.170.170/alaskadocs/Frontier_Travel_v_TSA_complaint.pdf.

[74] *Id.* at 4-5.

[75] *Id.* at 5.

[76] *Id.* at 8.

[77] *Id.* at 11.

Privacy Act of 1974, agencies must give actual notice that personal information is being collected, so that an individual can bring a timely claim challenging violations of his or her rights.[78] The plaintiffs requested the court to enjoin CAPPS II until the TSA and the DHS change their policies so they are in compliance with the Privacy Act. The plaintiffs also wanted notice provided when the agencies issue an order to implement the program.[79] The case is still pending in the court, and its status after the TSA's abandonment of CAPPS II is unknown.

### ii. REGISTERED TRAVELER PILOT PROGRAM

The TSA proposed the Registered Traveler Pilot Program (RT Pilot) on June 1, 2004, and issued its PIA on June 24, 2004, in order to assess a new type of passenger screening system mandated under the Aviation and Transportation Security Act.[80] The new pilot program was strictly voluntary, and allowed a registered traveler to provide personal information[81] which would be run through terrorist-related and criminal databases. If the passenger's information matches certain criteria set by the TSA, the information would be forwarded to the TSA for additional screening.[82] After the review, names of passengers considered to pose a security risk would be sent to law enforcement and/or intelligence agencies for detention or further investigation.[83] The TSA claimed that this layered procedural approach would prevent innocent people from being scrutinized at airports. The TSA also stated that it would use technical safeguards to prevent information from abuse and conduct privacy trainings for its agents, the DHS, and

---

[78] Brief for Frontier Travel at 10-11, *supra* note 73.

[79] *Id.* at 12.

[80] TRANSPORTATION SECURITY ADMINISTRATION, REGISTERED TRAVELER PILOT: PRIVACY IMPACT ASSESSMENT (Jun. 24, 2004), *available at* http://www.tsa.gov/interweb/assetlibrary/PIA_RT_OMB.pdf.

[81] It includes full name, social security number, other names used, home address, home telephone number, cell phone number, email address, date of birth, place of birth, nationality, gender, prior addresses for the past five years, drivers license number, and biometric identifiers (finger printing and/or iris scan). *Id.*

[82] *Id.*

[83] *Id.*

contractor staff who would have access to passenger information.[84] Because of the short duration of the program, however, the TSA makes clear that it does not provide any remedial measure for a person to correct erroneous information.[85] Despite these notifications, it is not clear if the TSA ever implemented this pilot program.

### iii. CREW VETTING PROGRAM

The TSA also announced the Crew Vetting Program and issued its PIA on June 28, 2004.[86] In October 2001, the Federal Aviation Administration issued an Emergency Amendment (EA) which required that countries highly concerned about security submit cockpit crew lists identifying names, dates of birth, places of birth, and pilot/flight engineering license numbers.[87] In December 2003, due to additional air security concerns, the TSA issued security directives and modified the EA by extending the scope to cover all flights flying over, in, and out of the United States, and to collect passport numbers in addition to those previously required.[88] On March 30, 2004, the TSA further modified the EA, by extending the coverage of the program to include cabin crews and persons on all cargo flights, and by requiring the agency to gather more information from these individuals, such as gender and job classification.[89]

The procedure for processing the crew information would be the same as it was for CAPPS II; it runs the information on various government databases – both non-classified and intelligence – to see if anyone matches with people of high security concern.[90] The PIA ensures that the information is used strictly for air safety purposes, even though it would be shared with other government agencies and intelligence, and that those who believe that the information the TSA

---

[84] *Id.*

[85] *Id.*

[86] TRANSPORTATION SECURITY ADMINISTRATION, PRIVACY IMPACT ASSESSMENT FOR CREW VETTING PROGRAM (Jul. 28, 2004), *available at* http://www.tsa.gov/interweb/assetlibrary/PIA_CVP_DHSCPO.pdf.

[87] *Id.* at 2.

[88] *Id.*

[89] *Id.*

[90] *Id.*

has on them is not correct can seek redress through U.S. embassies abroad.[91] The TSA had not determined, however, how long a record would be retained in its database.[92] So far, the extent to which the TSA has implemented this program is unknown to the public.

### iv. NO-FLY LIST

Aside from the CAPPS II, Pilot, and Crew Vetting Programs, the TSA maintains a No-Fly List, "a list circulated to commercial airlines and security personnel with instructions to detain and question any passenger whose name matches or is similar to one on the No-Fly List."[93] The List contains two different lists of individuals considered to be threats to air security: a no-fly list, which contains names of individuals who are prohibited from taking a flight, and a selectee list, which contains names of individuals who must go through additional security screening.[94] The TSA started implementing this program in November 2001.[95]

Some organizations, individuals, and members of Congress raised concerns over the No-Fly List. A class action was filed regarding the No-Fly List in the U.S. District Court for the Western District of Washington by the ACLU of Washington on April 6, 2004.[96] The Court heard the ACLU's argument on November 4, 2004.[97] This case involved seven plaintiffs and others similarly situated.[98] Upon check-in at airline counters, the plaintiffs were often informed by airline personnel that they would have to wait for long hours because there

---

[91] PRIVACY IMPACT ASSESSMENT FOR CREW VETTING PROGRAM, *supra* note 86, at 5-6.

[92] *Id.*

[93] Brief for Green at 1, *Green v. Transportation Security Administration* (D. W.Wash. 2004), *available at* http://www.aclu.org/SafeandFree/SafeandFree.cfm?ID=15419&c=272.

[94] *Id.* at 4.

[95] *Id.*

[96] *Id.* at 1.

[97] Press Release, American Civil Liberties Union, Federal Court Hears Arguments in ACLU's No-Fly List Case: ACLU Says Government List is Flawed and Violates Rights of Passengers (Nov. 4, 2004), *at* http://www.aclu-wa.org/Issues/otherissues/FedCrtNoFly.html.

[98] Brief for Green, *supra* note 93, at 1.

were problems in issuing boarding passes.[99] Some were informed that their names matched or were similar to the names listed on the No-Fly List.[100] The plaintiffs claimed that, as a result, they could not make scheduled flights and were inconvenienced, interrogated, and humiliated in front of other passengers, and received extra security checks at security checkpoints, even though they do not have any relation to terrorist organizations or have outstanding criminal records.[101] The plaintiffs argue that the TSA did not explain why they were identified by the No-Fly List.[102] In response, the plaintiffs have brought due process actions under the Fifth Amendment and privacy actions under the Fourth Amendment, seeking declaratory and injunctive relief.[103]

The No-Fly List received additional criticism from the public and Congress. At a hearing of the Senate Judiciary Committee in August 2004, Senator Ted Kennedy of Massachusetts stated that he had repeatedly been refused permission to board flights between Washington, D.C. and Boston in April of that year, because his name has been placed on the No-Fly List.[104] Because Senator Kennedy continued to have the problem even after the Secretary of the DHS, Tom Ridge, acknowledged the problem and apologized to him, Senator Kennedy was concerned about redress procedures for citizens who do not capture the attention of senior officials.[105] Georgia Representative John Lewis also testified that his name had been on the List.[106] The ACLU raised a similar concern about the No-Fly List, stating that there was "no effective mechanisms for removing names or identifying innocent passengers" and that not everyone was informed

---

[99] *Id.* at 3-24.

[100] *Id.*

[101] *Id.*

[102] *Id.*

[103] Brief for Green, *supra* note 93, at 26-27.

[104] Shaun Waterman, *Senator Gets a Taste of No-fly list Problem*, UNITED PRESS INT'L, Aug. 20, 2004 *available at* http://www.washingtontimes.com/upi-breaking/20040819-082444-1551r.htm.

[105] *Id.*

[106] Kehaulani Goo, *supra* note 29.

of the current TSA system of ombudsman for remedy.[107] In September 2004, 680 people filed claims for redress, and the TSA managed to clear only 250 by October.[108]

Also in 2004, the media reported several cases of inappropriate conduct by U.S. commercial airlines where they turned passenger information over to government agencies without any proper procedure. In February 2004, the DHS revealed that JetBlue Airways turned over PNRs, which were supposedly used for a purpose other than air safety, to a Department of Defense contractor.[109] American Airlines also admitted sending data to government contractors sometime after the terrorist attacks in 2001.[110] Similarly, Northwest Airlines acknowledged that it passed passenger information over to NASA in September 2001.[111]

## V. SECURE FLIGHT

Based on these criticisms, the TSA abandoned CAPPS II and introduced a new program called Secure Flight in September 2004. The goals of Secure Flight are: (1) to identify passengers known or suspected to be terrorists before flights; (2) to conduct efficient and quick passenger screening; (3) to reduce the number of passengers unnecessarily selected for secondary screening and better target known terrorists by excluding passengers without risk; and (4) to protect fully passengers' privacy and civil liberties.[112] In order to implement the program, the TSA mandated U.S.-based airlines to turn over PNRs

---

[107] Waterman, *supra* note 104.

[108] Christopher Elliott, *Getting Off a Security Watch List is the Hard Part*, THE N.Y. TIMES, Nov. 2, 2004, at C8, *available at* http://travel2.nytimes.com/mem/travel/article-page.html?res=9A0DE6D71E3DF931A35752C1A9629C8B63.

[109] Steven Roberts, *Big Brother is Watching You*, 27 NAT'L L. J. 15, (2004).

[110] Jay Boehmer, *TSA Demands PNR Data: Secure Flight Program Renews Privacy Issues*, BUSINESS TRAVEL NEWS, Oct. 4, 2004, *available at* http://www.btnmag.com/businesstravelnews/headlines/frontpage_display.jsp?vnu_content_id=1000651781.

[111] *Id.*

[112] Justin Oberman & Lisa Dean, Transportation Security Administration, Secure Flight, Presentation and Proposal to ASAC (Sept. 30, 2004). Powerpoint™ slides of the presentation are available at http://www.tsa.gov/interweb/assetlibrary/Secure_Flight_ASAC_Presenation.ppt#2.

from June 2004 by October 29, 2004.[113] This is estimated to include the records of fifty-four million Americans.[114] The TSA would compare airline PNR data to government and commercial data, and determine which components of the PNRs are necessary for pre-screening cross-checks once it completes testing.[115] TSA also began a 30-day period of soliciting comments from the public about Secure Flight on September 23, 2004.[116] The TSA assured that it would establish comprehensive passenger redress procedures, and personal data and civil liberties protections for the Secure Flight program.[117]

Secure Flight eliminated some of the controversial elements of CAPPS II. For instance, during the pre-screening phase under CAPPS II, the system could identify passengers with outstanding arrest warrants and refer them to law enforcement agents, whereas the Secure Flight system would only identify those suspected to be a threat to air flights.[118] Under Secure Flight, the TSA takes over the task of running airline passenger information with the "No-Fly List," which was previously conducted by airline employees.[119] The TSA stated that this procedure would allow the government to draw on a broader array of names of suspected terrorists from other intelligence agencies.[120]

The TSA published the PIA on the Secure Flight test phase on September 21, 2004.[121] The TSA identified several privacy issues and promised that it would fully protect passengers' privacy and civil

---

[113] Boehmer, *supra* note 110.

[114] Scott McMurren, *Give Thanks for Year's Blessings*, ANCHORAGE DAILY NEWS, Dec. 26, 2004, at G4.

[115] Oberman, *supra* note 112.

[116] Boehmer, *supra* note 110.

[117] Oberman, *supra* note 112.

[118] Roberts, *supra* note 109.

[119] Kehaulani Goo, *supra* note 29.

[120] *Id.*

[121] TRANSPORTATION SECURITY ADMINISTRATION, PRIVACY IMPACT ASSESSMENT; SECURE FLIGHT TEST PHASE, TSA-2004-19160 (Sept. 21, 2004) *available at* http://www.epic.org/privacy/airtravel/sf_pia_9.21.04.pdf.

liberties. In the use of commercial databases,[122] the TSA stated that it would test them only to identify instances in which passenger information is inaccurate or incorrect, and that it would not store the commercial data.[123] The test is strictly conducted "to ensure accuracy, efficacy and reliability."[124] Concerning notice or opportunities to consent, the TSA commented that airline passengers should be aware of certain things. The TSA said that by engaging in air travel, passengers have consented to certain screening protocols because passenger screening has already been in place and there are numerous media reports about the new program.[125]

According to the TSA, passenger information would be shared with TSA employees and contractors who have a "need to know" in order to conduct the required test, and would be used solely for the purpose of testing the Secure Flight program.[126] The security of databases would be safeguarded in accordance with the Federal Information Security Management Act of 2002[127] and also with policies and rules established by the TSA and the DHS.[128] Collected information will be retained by the TSA at the Office of National Risk Assessment in a secure facility for a period of time necessary to conduct and review the test.[129] In the near future, the TSA will issue a record retention schedule, which shows how much information is retained by the agency.[130] In addition, the agency will make sure that

---

[122] Commercial databases include those services to banking, home mortgage and credit industries. *Id.* at 5.

[123] *Id.* at 5-6.

[124] *Id.* at 4.

[125] *Id.* at 6.

[126] *Id.* at 7.

[127] Pub. L. No. 107-347. (This act "established government-wide computer security and training standards for all persons associated with the management and operation of Federal computer systems." TRANSPORTATION SECURITY ADMINISTRATION, *supra* note 121, at 7.)

[128] Such rules include password protection and firewall protection. TRANSPORTATION SECURITY ADMINISTRATION, *supra* note 121, at 7-8.

[129] TRANSPORTATION SECURITY ADMINISTRATION, PRIVACY ACT OF 1974: SYSTEM OF RECORDS; SECURE FLIGHT TEST RECORDS, *TSA-200419160*, *available at* http://www.tsa.gov/interweb/assetlibrary/Secure_Flight_SORN_9.21.04.pdf.

[130] *Id.*

a proper redress procedure is in place and that passengers have access to their own PNRs.[131]

Along with the PIA, the TSA gave a thirty-day period for the public to comment on Secure Flight.[132] On November 10, 2004, the TSA issued a report titled "Notice of Final Order for Secure Flight Test Phase and Response to Public Comments on Proposed Order."[133] In this notice, the TSA responded to the public comments submitted during the thirty-day period. In the report, the following eight issues of public concerns regarding Secure Flight were identified: (1) the program's effects on privacy and civil liberties; (2) the routine use of information; (3) passenger consent to the use of PNRs; (4) the absence of redress mechanisms; (5) concerns over the use of commercial databases; (6) the efficacy of the program; (7) TSA's compliance with the Privacy Act of 1979 and other laws, regulations, and rules; and (8) potential conflicts with the European Union's privacy laws.[134]

The Final Order answered those concerns raised by the public. The National Business Travel Association (NBTA) stated that the TSA should balance the need to establish better security with policies and procedures that protect civil liberties and privacy.[135] Also, the Electronic Privacy Information Center (EPIC) raised a concern over the TSA's previous statement that the Secure Flight test phase should be exempted from the provisions of the Privacy Act.[136] An example of such a provision is the right of individuals to access government information about them.[137] The TSA justified its actions by stating that the Privacy Act specifically allowed agencies to exempt information regarding national security or law enforcement concerns from the Act's authority.[138] In answering the eight questions, the TSA

---

[131] *Id.*

[132] *Id.*

[133] TRANSPORTATION SECURITY ADMINISTRATION, NOTICE OF FINAL ORDER FOR SECURE FLIGHT TEST PHASE; RESPONSE TO PUBLIC COMMENTS ON PROPOSED ORDER AND SECURE FLIGHT TEST RECORDS, TSA-2004-19160 (Nov. 10, 2004), *available at* http://www.tsa.gov/interweb/assetlibrary/secureflightfinalorder.pdf.

[134] *Id.* at 7.

[135] *Id.* at 8.

[136] *Id.* at 9.

[137] *Id.*

[138] *Id.*

mostly repeated its PIA by saying that it will take steps to ensure privacy, and described the same precautions and measure the agency said it would take.[139]

As a result, the Final Order only has a few changes from its original plans announced in September.[140] The program applies only to flight segments completed in June 2004, instead of PNRs with flight segments before June 2004.[141] The order also clarifies that the program applies to public charter flights, while international flights to and from the U.S. are completely excluded from the program.[142]

The TSA estimates that testing of the new program will continue through spring to early summer 2005 when the first airlines are expected to begin implementing the system.[143] After the test phase is completed, if the TSA implements Secure Flight as a permanent program, it must face the GAO's objective evaluation.[144] This process is required by law and the TSA cannot implement its plan until after GAO's approval.[145]

Secure Flight still faces some criticism from the public. For instance, the Business Travel Coalition, which represents American businesses, commented that the new program is another CAPPS II and just a "replacement for the original CAPPS."[146]

## GOVERNMENT DATA MINING

Government agencies and departments gather information for various purposes. The collection of data is useless unless the data has meaning, however. The government and various private entities are now taking data mining a step further, by gathering various data into one source and then analyzing it to obtain new data. This is the

---

[139] *Id.* at 7-26.

[140] *Id.* at 26.

[141] *Id.* at 27.

[142] *Id.*

[143] Kerry Ezard, *First US Carriers to Implement Secure Flight in Late Spring*, AIR TRANSP. INTELLIGENCE, Dec. 1, 2004.

[144] Boehmer, *supra* note 110.

[145] *Id.*

[146] *Id.*

process known as data mining. This process is being used commercially to assist businesses in targeting consumers and improving processes, as well as by state and federal governments for a variety of purposes. The government's use of data mining programs requires a balance between privacy interests and effectively using available data for useful purposes.

Privacy concerns have been raised with regard to several federal government initiatives, such as the defunct Total Information Awareness Program and the Intelligence Reform and Terrorism Prevention Act of 2004,[147] as well as state data mining projects like the Multistate Anti-Terrorism Information Exchange (MATRIX). Discussions are currently underway to determine how government data mining can be limited to protect privacy, while maintaining or enhancing its usefulness. The use of private institutions as a source for government data mining is also a topic of current debate. The following analysis deals with these federal and state programs, as well as general debates and controversies that were prominent throughout 2004.

## I. DATA MINING OVERVIEW

It is important to understand the nature of data mining, in both the private and public sectors, before one can evaluate federal and state data mining programs and the privacy concerns that arise from the use of those programs. The following section gives a broad overview of data mining by analyzing how it is defined, the purpose of the technique, and the extent to which the government uses it. The last component focuses on the concerns and challenges that arise from government data mining.

### A. DATA MINING DEFINED

Data mining has been defined to incorporate many parameters of the process, including technology, analysis, extraction, and knowledge. The General Accounting Office defines data mining as the "application of database technology and techniques – such as statistical analysis and modeling – to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the

---

[147] Intelligence Reform and Terrorism Prevention Act, Pub. L. No. 108-458 (codified at 50 U.S.C. 401).

prediction of future results."[148] Simpler definitions focus on the fact that data mining is a "technique for extracting knowledge from large volumes of data"[149] or that the process looks for "trends or anomalies" in the data.[150] With regard to recent government uses for data mining, the technique can be defined as the use of "computer technologies to sift through large data repositories to identify threatening patterns and people" in an effort to predict crime or terrorist activity prospectively.[151]

Data mining has a variety of uses. Many consider it to be a "sense-making application" because it brings meaning to large conglomerations of raw data that previously did not make sense separately.[152] The information that is used is known information that has generally been previously disclosed, and may be an accumulation of information from a variety of sources. Data mining then develops information that was previously unknown by the system out of the known data.[153] This unknown information that is generated may be a hidden pattern that can be used to predict future behavior.[154] These predictions prove to be valuable for commercial uses, such as developing marketing strategy, or for public use.

Data mining was originally a product of IBM.[155] However, it is currently developed and used by numerous private and public entities. Not only does the government develop its own data mining processes, but it also purchases this technology from other private or educational

---

[148] U.S.GENERAL ACCOUNTING OFFICE, DATA MINING: FEDERAL EFFORTS COVER A WIDE RANGE OF USES1 GAO-04-548 (May 2004), *available at* http://www.gao.gov/new.items/d04548.pdf.

[149] Eric J. Sinrod, *What's Up with Government Data Mining?*, USA TODAY (June 9, 2004), *available at* http://www.usatoday.com/tech/columnist/ericjsinrod/2004-06-09-sinrod_x.htm.

[150] SOURCEWATCH, *Data Mining*, *at* http://www.sourcewatch.org/wiki.phtml?title=Data_mining (last modified May 27, 2004).

[151] Anita Ramasastry, *The Safeguards Needed for Government Data Mining*, FINDLAW (Jan. 7, 2004), *at* http://writ.news.findlaw.com/ramasastry/20040107.html.

[152] K.A. Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 COLUM. SCI. & TECH. L. REV. 2, 6 (2003/2004).

[153] *Id.*

[154] SOURCEWATCH, *supra* note 150.

[155] *Id.*

institutions.[156] For example, ChoicePoint is one of the largest suppliers of data to the government.[157] This company has nearly fourteen billion records from which it mines data, and makes the data available to the government.[158] Data mining usually involves computer software that "extracts information from databases, as well as text, voices, other audio, video, graphs, images, maps, and equations and chemical formulas."[159] It then uses this information to search using varying parameters such as sequence, clustering, association, classification, and forecasting to determine information about a specific subject or to determine a pattern.[160]

## B. PURPOSE OF DATA MINING

Government data mining is used for a variety of purposes throughout the federal and state structures. The most recognized use of government data mining is the analysis of intelligence to detect terrorist or criminal activities or patterns.[161] Although this is a purpose served through government data mining, the reach of the process is much wider. The most common use of data mining cited by government agencies was to improve service or performance.[162] Government agencies also specified other uses, such as managing human resources activities or patterns, detecting and preventing fraud, and generating statistics to be used in audits or investigations.[163]

---

[156] Associated Press, *Controversial Government Data Mining Research Lives On* (Feb. 23, 2004), THE MERCURY NEWS *at* http://www.siliconvalley.com/mld/siliconvalley/news/editorial/8022436.htm.

[157] Peter Jennings, *No Place to Hide: Freedom and Identity*, (Jan. 20, 2005), *at* http://abcnews.go.com/Technology/Primetime/story?id=429308&page=1.

[158] Lee Tien, *Privacy, Technology and Data Mining*, 30 OHIO N.U.L. REV. 389, 389 (2004).

[159] Associated Press, *supra* note 156.

[160] SOURCEWATCH, *supra* note 150; Jeffrey W. Seifert, *Data Mining: An Overview*, CONGRESSIONAL RESEARCH SERVICE 1 (Dec. 16, 2004).

[161] Thomas Claburn, *GAO Raises Privacy Concerns about Federal Data Mining*, INFORMATIONWEEK (June 4, 2004), *at* http://informationweek.com/story/showArticle.jhtml?articleID=21401674.

[162] *Id.*; U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 2.

[163] U.S.GENERAL ACCOUNTING OFFICE, *supra* note 148, at 2,3; Claburn, *supra* note 161.

Many programs for government data mining are conducted for research and development purposes. These programs will later be used by various government and private institutions but are currently being refined to enhance their usefulness. Many believe that the technology will be developed, with or without government involvement; therefore, it is better for the government to develop it so there are checks on the potential invasion of privacy from the technology.[164] The argument rests on the idea that the government is generally more accountable to the public than private institutions. Others counter this argument by undermining the basis of the argument. They believe that the private sector is subject to stricter privacy standards than the government.[165] For example, the private sector has significant restrictions on the release of personal financial data under the Fair Credit Reporting Act, which the government is not subject to.[166]

### C. USAGE OF GOVERNMENT DATA MINING

After September 11, 2001, the United States government was chastised for its inability to "connect the dots" and prevent the terrorist attacks. Due to this criticism, the government made updating information technology a priority, and "information sharing and automated analysis technologies have become part of official government information technology development policy."[167] A May 2004 report by the General Accounting Office revealed the extent that the federal government uses data mining in their operations.[168] The report indicates that 52 out of 128 federal agencies, nearly forty percent, participate in some level of government data mining.[169] Among these 52 agencies, there are a total of 199 data mining programs, with 131 of these in progress and 68 planned.[170] The

---

[164] Taipale, *supra* note 152, at 5.

[165] Robert Pear, *Panel Urges New Protection on Federal 'Data Mining,'* N.Y.TIMES, May 17, 2004, at A12.

[166] *Id.*; *see* Fair Credit Reporting Act, 15 U.S.C. §1681.

[167] Taipale, *supra* note 152, at 2.

[168] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148.

[169] *Id.* at 2; Claburn, *supra* note 161.

[170] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 3; Claburn, *supra* note 161.

Department of Defense currently uses 47 data mining projects, employing the largest number of projects.[171] Out of the 199 projects being used, 77 of them have participated in data sharing.[172] They use data obtained from other federal agencies, as opposed to data original to that agency or obtained from a private entity. Additionally, 122 of the programs used personal information, such as phone numbers, social security numbers, email addresses, and driver's license numbers.[173] The potential for future data mining is exponential. There are currently over 2,000 databases within federal agencies and departments.[174] Many of these could be subject to data mining within that agency or shared with other agencies for the purpose of data mining.

### D. CHALLENGES REGARDING GOVERNMENT DATA MINING

When the government mines data, great amounts of useful information may be found. There are, however, significant challenges that are faced when participating in this process. First, there are concerns over the quality and accuracy of the data being mined.[175] Accurate information is vital to producing useful results. Without current and accurate information more mistakes will arise, such as mistaken identity and the misinterpretation of information leading to faulty inferences. The quality of the data may be especially dependent upon the source of the data. Information gathered by the private sector then obtained by the government is the most likely set of data to be inaccurate.[176] For example, it is more likely that one would report a false address when registering at a hotel than when that person is obtaining his or her driver's license from the government. It is

---

[171] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 3; *See* Kim Zetter, *GAO: Fed Data Mining Extensive* (May 2, 2004), *at* http://www.wired.com/news/privacy/0,1848,63623,00.html?tw=wn_story_related; Pear, *supra* note 165, at 12.

[172] SOURCEWATCH, *supra* note 150.

[173] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 10; *See* Pear, *supra* note 165, at 12.

[174] Tien, *supra* note 158, at 389.

[175] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 6; Seifert, *supra* note 160, at 11.

[176] James X. Dempsey, The Defense of Privacy Act and Privacy in the Hands of The Government, Statement before the House Committee on the Judiciary Subcommittee on Commercial and Administrative Law and Subcommittee on the Constitution (July 22, 2003), *available at* http://judiciary.house.gov/HearingTestimony.aspx?ID=150.

suggested that "data cleansing" is a key requirement "to achieving useful results."[177] Data cleansing requires the information to be sorted through and checked for accuracy. This can be a daunting task considering that vast amounts of data are available.[178]

Some argue, however, that the government's reliance on the information produced through data mining should be curbed, unless there are data quality or reliability standards in place.[179] Without these standards, the information should not be used to generate probable cause, and should be limited to finding data quickly if there is already particularized suspicion concerning an individual.[180] Additionally, the quality of the information may depend on the skill of the analyst.[181] Data mining systems are complex and require a highly trained analyst to bring value to the information produced. If the skill of the analyst is lacking, then the quality of the information produced could be lacking as well.

Third, there is a concern that the lack of procedural, judicial, and congressional constraints on data mining results in increased privacy violations. There are currently a limited number of judicial or procedural limits placed on both commercial and government programs.[182] It is argued that without limits, the security of the data can be easily compromised.[183] For example, there may be intentional abuses by government employees that have access to the data, even though their access is not necessary for the effectiveness of the

---

[177] Taipale, *supra* note 152, at 7.

[178] Letter from U.S. Senator Patrick Leahy to The Honorable John Ashcroft, Attorney General 2 (Jan. 10, 2003), *available at* http://www.fas.org/sgp/news/2003/01/leahy011003.html.

[179] James X. Dempsey & Paul Rosenzweig, *Technologies that Can Protect Privacy as Information is Shared to Combat Terrorism,* CENTER FOR DEMOCRACY & TECHNOLOGY 4 (May 26, 2004), *available at* http://www.cdt.org/security/usapatriot/20040526technologies.pdf.

[180] *Id.* at 5; Taipale, *supra* note 152, at 18.

[181] Seifert, *supra* note 160, at 11.

[182] Ramasastry, *supra* note 151.

[183] Gregory D. Kutz, Data Mining: Results and Challenges for Government Program Audits and Investigations, Testimony before the Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census, Committee on Government Reform, House of Representatives, U.S. GENERAL ACCOUNTING OFFICE, GAO-03-591T (Mar. 25, 2003), *available at* http://www.gao.gov/new.items/d03591t.pdf.

process.[184] In the commercial sector, the security of a ChoicePoint database was compromised, allowing for the identity theft of 145,000 individuals across the United States.[185] Judicial and procedural constraints could help answer questions such as: Who can access data? What kind of data can and will be compiled? How will data sharing be enforced? How do we guarantee that those who do not have access will not be able to access the data?[186] Rules that govern how technology operates, rather than the type of technology used and developed, can be used to establish privacy standards.[187]

Fourth, there is a general concern regarding the use of data for purposes other than that for which it was originally collected.[188] When data is shared and mined or mined for various reasons by the collector, the information is being used in a way that may not have been anticipated by the holder of the information when the information was given. For example, there is concern that data will be gathered and used in the name of fighting terrorism but will actually be used for general law enforcement.[189] Some privacy advocates and civil libertarians believe that consent should be required before information can be used for a purpose other than that for which it was gathered. Others would agree that rules allowing an individual to know about the "collection of their personal information, how to access that information, and how to request a correction of inaccurate information" should be considered by lawmakers.[190] Such rules would require a balance between citizen consent and maintaining the usefulness of the program.

Fifth, numerous Fourth Amendment issues arise through the use of data mining. The basic purpose of the Fourth Amendment "is to safeguard the privacy and security of individuals against arbitrary invasions by government officials."[191] Although many may think of

---

[184] Dempsey & Rosenzweig, *supra* note 179, at 4.

[185] Bob Sullivan, *Data Theft Affects 145,000 Nationwide,* MSNBC (Feb. 18, 2005), *available at* http://www.msnbc.msn.com/id/6979897/.

[186] Ramasastry, *supra* note 151.

[187] Taipale, *supra* note 152, at 3.

[188] Seifert, *supra* note 160, at 12.

[189] Dempsey & Rosenzweig, *supra* note 179, at 5.

[190] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 6.

[191] Tien, *supra* note 158, at 400.

this principle applying to law enforcement specifically, there is no categorical exclusion of other government officials. This implies that a violation of privacy by an official in the Department of the Interior would invoke Fourth Amendment protection.

Two constitutional privacy concerns arise from data mining. First, the core of the privacy complaint is that data is being linked.[192] If the information was initially given voluntarily, then a privacy violation is doubtful. However, by linking this voluntary information, one can determine previously unknown information that would have required consent to obtain if it were obtained without the data mining process. Also, associational privacy may be violated.[193] Data mining searches patterns and determines a person's relationships and behaviors, so that it can determine associations. A typical example is the use of data mining to determine an individual's association to a terrorist organization.

The Fourth Amendment does not apply unless the data mining is considered a search.[194] In order for an activity to be a search, the individual must have a reasonable expectation of privacy.[195] There are several arguments both identifying data mining as a search or not as a search. Those that argue that there is no reasonable expectation of privacy in these documents, and that data mining therefore is not a search, base their argument on the fact that the information is knowingly exposed. The information was knowingly exposed to the public when the individual released it, and knowingly exposed information has no reasonable expectation of privacy. Therefore, the data mining is not a search.[196]

Alternatively, those believing data mining is a search argue that the mere fact that the information has been exposed does not make the Fourth Amendment irrelevant.[197] They argue that although the original information was knowingly exposed, the patterns and behaviors that are discovered were not, and the individual holds an expectation of privacy concerning them. Those advocating for Fourth

---

[192] *Id.* at 398.

[193] *Id.* at 399.

[194] *Id.* at 408.

[195] *Id.*

[196] Tien, *supra* note 158, at 408.

[197] *Id.* at 393.

Amendment protections in data mining believe that federal agencies should obtain court approval before engaging in data mining, especially data mining that deals with personally identifiable information.[198] This court approval would be based on particularized suspicion.[199] Basing the data mining on particularized suspicion would limit inquiries concerning an individual, unless there was already existing suspicion as to illegal activities. The court approval would also be based on particularized scope. For example, the generalized "fishing expeditions" that are used by many data mining programs would be a violation of the Fourth Amendment.[200]

## II: FEDERAL DEVELOPMENTS IN GOVERNMENT DATA MINING

Federal agencies are the primary users of public sector data mining initiatives. During 2003, significant dilemmas in data mining were placed in the public eye and the government was forced to balance the privacy interests against the necessity and usefulness of certain data mining programs. One program that did not withstand this balancing test was the Total Information Awareness Program, which received a funding cut in September 2003. This program, along with the 2004 report from the National Commission on Terrorist Attacks upon the United States (9/11 Commission) encouraged Congress to look at how the intelligence community shares and utilizes information, which resulted in the passage of the Intelligence Reform and Terrorism Prevention Act of 2004. The following section looks at these developments along with other data mining initiatives that developed in 2004.

### A. TOTAL INFORMATION AWARENESS PROGRAM

The Total Information Awareness Program (TIA) was one of the most publicized and controversial data mining projects conducted by the United States government in recent years. This use of technology and research generated much criticism, and ultimately funding for the program was eliminated by Congress. Parts of the program continue in various forms, however.

---

[198] *Id.* at 402.

[199] *Id.*

[200] *Id.* at 405.

### i. TOTAL INFORMATION AWARENESS PROGRAM OVERVIEW

The TIA was developed as a project of the Defense Advanced Research Projects Agency under the direction of John Poindexter, who was appointed head of the Information Awareness Office to lead this project.[201] The TIA was designed to aggregate and analyze information from a wide array of databases.[202] The program would then create a baseline pattern identifying suspicious behavior and mine the data based on that pattern.[203] Poindexter expressed that the purpose of the TIA was solely for research and development of the technology and to create a prototype, not for implementation.[204] Three areas were identified as technologies to be developed through the TIA: language translation capacity, data search with pattern recognition and privacy protection, and advanced collaborative and decision support tools.[205] Congress would be responsible for determining whether it violated privacy rights, and the extent to which the technology would be used.[206]

The development of the TIA was spurred by the September 11, 2001 terrorist attacks. The theory behind its development was that possible terrorist threats could be identified by sorting through everyday transactions, such as credit card purchases, car rentals, and travel reservations.[207] The development of the TIA technology was to be used to produce watch lists, profiles, and to mark certain

---

[201] PEOPLE FOR THE AMERICAN WAY, DATA MINING (Jan. 7, 2003), *available at* http://www.pfaw.org/pfaw/general/default.aspx?oid=9718&print=yes&units=all; Interview by Robert O'Harrow with John Poindexter, Head of the Total Information Awareness Program, Defense Advanced Research Projects Agency (2004), *available at* http://www.noplacetohide.net/poindexter.html.

[202] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[203] Sharon R. Anderson, *Total Information Awareness and Beyond: The Dangers of Using Data Mining Technology to Prevent Terrorism,* BILL OF RIGHTS DEFENSE COMMITTEE, *available at* http://www.bordc.org/data-mining.pdf; Taipale, *supra* note 152, at 11.

[204] Paul Rosenzweig, *Proposals for Implementing the Terrorism Information Awareness System,* THE HERITAGE FOUNDATION (Aug. 7, 2003), *available at* http://www.heritage.org/Research/HomelandDefense/lm8.cfm.

[205] Seifert, *supra* note 160, at 5.

[206] Interview by Robert O'Harrow with John Poindexter, *supra* note 201.

[207] *Id.*

individuals as suspicious.[208] This was in an effort to prevent future terrorist attacks, primarily by analyzing patterns of activity, rather than individual activity.[209]

### ii. CRITIQUE OF THE TOTAL INFORMATION AWARENESS PROGRAM

Many of the concerns of data mining in general have been expressed toward the TIA, but with greater specificity. Privacy advocates felt that the TIA presented an "enormous invasion of privacy."[210] Even the Department of Defense Inspector General stated that the program failed to meet privacy concerns.[211] There was a sense that the data mining was delving into areas of information that pushed the envelope of Fourth Amendment protections. "Data mining, like any other government data analysis, should occur where there is a focused and demonstrable need to know, balanced against the dangers to civil liberties. It should be purposeful and responsible."[212] Some believe that the TIA constituted domestic spying and was intended solely for surveillance of Americans.[213] This became the fear of many Americans based on the name, the logo which resembled an omnipotent, watchful eye, and the motto that "knowledge is power."[214] There was apprehension that the government, through the TIA, was becoming a big brother and had catalogues of information about the private lives of all citizens.[215]

A Fourth Amendment debate over the program ensued over what level of suspicion was currently required and should be required to

---

[208] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[209] Associated Press, *supra* note 156; Rosenzweig, *supra* note 204.

[210] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[211] *Id.*

[212] Letter from U.S. Senator Patrick Leahy to The Honorable John Ashcroft, *supra* note 178.

[213] Electronic Privacy Information Center, *Poindexter's Recent Op-Ed Reflects Inconsistencies in Statements Regarding Total Information Awareness (TIA)* (2004) *available at* http://www.epic.org/privacy/profiling/tia/TIAinconsistencies.html.

[214] Seifert, *supra* note 160, at 7.

[215] American Civil Liberties Union, *Stop the Government Plan to Mine Our Privacy with "Total Information Awareness" System* (2003), *available at* http://www.aclu.org/Privacy/Privacy.cfm?ID=11323&c=130; Rosenzweig, *supra* note 204.

justify the mining of personal information. There was an argument that the program worked backwards and would make all citizens suspects of terrorist activity without proof of any wrongdoing.[216] Rather than being used to generate probable cause, the TIA should only be employed after probable cause has already been generated.[217] Accessing sensitive data to determine if someone could be a terrorist without any probable cause is a violation of the Fourth Amendment.[218] Another argument against the TIA's violation of the Fourth Amendment was based on search warrants. These arguments assumed that individuals had an expectation of privacy with regard to their data, and therefore TIA data mining constituted a search. However, there was no probable cause, particularized suspicion, or warrant obtained so the searches were unreasonable and did not comply with the Fourth Amendment.[219] The only way that the warrantless search could fall within the Fourth Amendment would be if it fit within one of the exceptions of incident to arrest, exigent circumstances, or consent.[220] Most cases do not fit under these categories and therefore, it is argued, would be unreasonable and a violation of privacy rights. Regardless of the theory used to determine an intrusion on Fourth Amendment privacy rights, it was generally agreed that the TIA did not fully protect the privacy of American citizens.

An additional critique of the program was its lack of safeguards. With the enormous size of the database and the wealth of information that could be gathered from the system, the safeguards put in place were inadequate.[221] Arguments were made to Congress that the TIA database was a prime target for "exploitation and attack by malicious computer users."[222] Not only was the database subject to outside

---

[216] American Civil Liberties Union, *supra* note 215.

[217] Associated Press, *Give It Up: Info for Protection,* WIRED.COM (May, 2, 2004), *available at* http://www.wired.com/news/privacy/0,1848,63304,00.html?tw=wn_story_related.

[218] Ramasastry, *supra* note 151.

[219] Max Blumenthal, *Data Debase*, PROSPECT.ORG (Dec. 19, 2003), *available at* http://www.prospect.org/webfeatures/2003/12/blumenthal-m-12-19.html; Anderson, *supra* note 203, at 12.

[220] Anderson, *supra* note 203, at 13.

[221] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[222] Letter from the Association for Computing Machinery's U.S. Public Policy Committee to Senators John Warner & Carl Levin 1 (Jan. 23, 2003), *available at* http://www.acm.org/usacm/Letters/tia_final.html.

attackers, but also to abuse by internal users. Thousands of users had and thousands more would have access to this system, which would make security difficult and the potential for extortion high.[223] These lack of safeguards could lead to data being exploited, therefore resulting in even further invasions of privacy than what the TIA was already accused of.[224]

As with data mining in general, many perceived the TIA as an inaccurate source of information. There may be difficulties in continually having updated or accurate information considering that the database has input from a variety of sources.[225] This could result in a high number of false positives, causing innocent people to be placed on watch lists or preventing them from obtaining certain jobs.[226] Another factor that may dilute the effectiveness of the program is acclimation. The targets, such as potential terrorists, will quickly learn the parameters and patterns being used to identify them and quickly change their habits.[227] Without skilled analysts and technology support staff to make updates to the program continuously, it could quickly become outdated and ineffective.

### iii. MORATORIUM ON THE TOTAL INFORMATION AWARENESS PROGRAM

If the TIA were completed, it would have been the largest domestic surveillance system in the United States.[228] However, the overwhelming concerns with the program prompted Congress to place a moratorium on funding for the program in September 2003. This moratorium was part of a FY2004 Department of Defense Appropriations Act.[229] It stated that there would be no further

---

[223] *Id.*; Declan McCullagh, *TIA Proponents Defend Domestic Spy Plan,* CNET (Apr. 2, 2003), *available at* http://news.com.com/TIA+proponents+defend+domestic+spy_plan/2100-1029_3-995229.html?tag=nl.

[224] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[225] *Id.*

[226] Letter from U.S. Senator Patrick Leahy to The Honorable John Ashcroft, *supra* note 178; Associated Press, *supra* note 156; PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[227] Anderson, *supra* note 203, at 12.

[228] Blumenthal, *supra* note 219.

[229] Seifert, *supra* note 160, at 7; Data Mining Moratorium Act of 2003, Pub. L. No. 108-87.

appropriations to the TIA or other similar programs.[230] The two primary reasons that Congress suspended funding were that (1) there was no evidence that data mining was an effective tool for preventing terrorism and (2) privacy rights could be adversely affected because private information was mined.[231] The Act also required a report on the current state of data mining to be delivered to Congress.[232] This report was generated and delivered in the form of the GAO 2004 report.[233]

### iv. THE FUTURE OF THE TOTAL INFORMATION AWARENESS PROGRAM OR SIMILAR PROGRAMS

The future of the TIA is uncertain. It is clear that Congress will not support funding for the TIA in its current state. However, various groups have suggested improvements to the TIA that may make it useful while alleviating privacy invasion concerns. The Heritage Foundation has suggested the following safeguards that may achieve these goals:[234]

- Congressional authorization for data mining.
- Built-in limitations on access to third party data. For example, a subpoena could be required to search credit card or bank records.
- Patterns to identify terrorists to be approved by a Senate-appointed official before the search is conducted.
- Individual information to be disaggregated from pattern analysis.

---

[230] The Data-Mining Moratorium Act of 2003 (Jan. 16, 2003), http://www.techlawjournal.com/cong108/datamining/20030116.asp.

[231] *Id.*

[232] *Id.*

[233] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148.

[234] Rosenzweig, *supra* note 204.

- Individual identities not to be disclosed without approval from a federal judge.
- Data mining leading only to further investigation, not arrest or a no fly list.
- Implement a system to correct mistakes and false positives.[235]
- Additional accountability and oversight by other government branches.
- Restriction that TIA can only be used for terrorism investigations.
- Allow for tort remedies or administrative hearings for individuals harmed due to wrongful identification.[236]

Despite the moratorium ending all funding, certain research and development aspects of the TIA have continued under other programs and agencies. Essentially some of the technology that was being developed under the name of the TIA was transferred to other intelligence offices to continue to be developed.[237] For example, the voice recognition software that the TIA began constructing has been moved to the research and development office at the Pentagon.[238] Another project that has continued is the creation of translation software that mines data through spoken archives.[239]

Although the moratorium restricted funding for programs similar to the TIA, several have remained. One of the most prominent is the National Foreign Intelligence Program.[240] This program is a joint

---

[235] *See also id.*

[236] *See also id.*

[237] Associated Press, *supra* note 156.

[238] SOURCEWATCH, *Total Information Awareness* (2004), *available at* http://www.sourcewatch.org/wiki.phtml?title=Total_Information_Awareness.

[239] Blumenthal, *supra* note 219.

[240] Associated Press, *supra* note 156; CENTER FOR DEMOCRACY AND TECHNOLOGY, *Fact Sheet: Data Mining Programs and Other Government Uses of Commercial Data* (Oct. 16, 2003), *available at* http://www.cdt.org/security/usapatriot/031016factsheet.shtml.

collaboration between the Central Intelligence Agency (CIA), the Federal Bureau of Investigation (FBI), and the National Security Agency.[241] It does not invoke the privacy issues that the TIA did, because it can only be used overseas or against non-U.S. citizens in the United States. It cannot be used against Americans on American soil.[242] This is the primary difference between this permitted data mining program and the stricken TIA. Another similar program that did not receive a funding cut is a project of the Advanced Research and Development Activity.[243] This $64 million project has similar data mining feature to that of the TIA, and even uses former TIA researchers and analysts.[244]

With the news that a number of TIA projects are continuing in various departments, privacy advocates worry that TIA projects that the public does not know about are continuing to be developed.[245] Essentially, the structure of the TIA may be dissolved, but the components that presented privacy issues are still vibrant. To combat the implementation of these programs, several suggestions have been made. First, privacy impact assessments should be made for each government data mining project.[246] These PIAs could be conducted before research and development begins, at the prototype stage, and immediately prior to implementation. Another suggestion would be to have a Privacy Ombudsman oversee the development and implementation of government data mining projects.[247]

### B. INTELLIGENCE REFORM AND TERRORISM PREVENTION ACT OF 2004

In December 2004, Congress passed and the President signed into law the Intelligence Reform and Terrorism Prevention Act of 2004 (IRTPA). The purpose of this act is "to assist in connecting the dots of

---

[241] Blumenthal, *supra* note 219.

[242] Interview by Robert O'Harrow with John Poindexter, *supra* note 201.

[243] Associated Press, *supra* note 156.

[244] *Id.*

[245] Blumenthal, *supra* note 219; Anderson, *supra* note 203, at 4.

[246] Ramasastry, *supra* note 151.

[247] *Id.*

intelligence information" particularly in regard to terrorist activity.[248] The text of the bill makes clear that information sharing will be encouraged among the intelligence community. Many Congresspersons suggested similar bills that would require uniform intelligence sharing and the creation of a centralized director of intelligence. For example, the Shield Privacy Act was proposed in May 2004 and would create a Privacy Czar in the Office of Management and Budget.[249] However, the ideas that won the vote of both Houses and the President came in the summer of 2004 from the National Commission on Terrorist Attacks Upon the United States (9/11 Commission).[250] This commission criticized the intelligence community for its fragmented management structure and inability to "connect the dots" to pre-empt terrorist strikes.[251] The act consists of various provisions such as increased transportation and border security and terrorism prevention mechanisms;[252] however, this article will only focus on the parts that may affect data mining and how privacy may be affected by the new act.

### i. CREATION OF THE DIRECTOR OF NATIONAL INTELLIGENCE

One of the most dramatic changes made by the IRTPA was the addition of a Director of National Intelligence (DNI). The purpose of the DNI varies depending on the source, but Senator Bob Graham endorsed the bill by stating that, "the intelligence community needs a leader with the clout to set common goals, establish priorities, knock heads, and ensure that the American people are protected."[253] There is

---

[248] Rosenzweig, *supra* note 204.

[249] Kim Zetter, *U.S. May Get a Privacy Czar,* WIRED.COM (May 21, 2004), *at* http://www.wired.com/news/privacy/0,1848,63542,00.html?tw=wn_story_related.

[250] Todd B. Tatelman, *Intelligence Reform and Terrorism Prevention Act of 2004: National Standards for Drivers' Licenses, Social Security Cards, and Birth Certificates,* CONGRESSIONAL RESEARCH SERVICE 1 (Dec. 16, 2004), *available at* http://www.fas.org/irp/crs/RL32722.pdf.

[251] Alfred Cumming, *The Position of Director of National Intelligence: Issues for Congress,* CONGRESSIONAL RESEARCH SERVICE 1 (Aug. 12, 2004) *available at* http://www.fas.org/irp/crs/RL32506.pdf.

[252] U.S. EMBASSY OF CANADA, Intelligence Reform and Terrorism Prevention Act of 2004 (2005), *available at* http://www.usembassycanada.gov/content/can_usa/borderissues_irtpa2004.pdf.

[253] Cumming, *supra* note 251, at 14.

general agreement that the DNI will create one entity through which intelligence budget, personnel, and information will flow.[254] Other stated purposes have been to ensure privacy protections are not eroded by intelligence technologies, to coordinate all intelligence efforts of the federal government, and to lighten the burden of the Director of Central Intelligence.[255]

Congress has made at least fourteen attempts to create a Director of National Intelligence or similar position.[256] Between June 2002 and June 2004, at least six bills were introduced in the Senate in an attempt to create the position.[257] In 2002, Senator Dianne Feinstein introduced a proposal that was rejected, but its revised version, which gave the DNI budget and personnel responsibilities, finally became the IRTPA.[258]

Below is a list of intelligence and privacy related responsibilities that will be taken on by the DNI.

- Report directly to the President and provide daily briefings to the President about intelligence developments.[259]

- Establish objectives and priorities for the intelligence community.[260] Privacy advocates are pushing for one of these objectives to be a protection of individual privacy rights.

- Determine and execute a budget for the National Intelligence Program and has the ability to transfer funds

---

[254] *Id.* at 5.

[255] *Id.* at 13.

[256] *Id.* at 16–19.

[257] *Id.* at 4, 5.

[258] Cumming, *supra* note 251, at 4.

[259] U.S. EMBASSY OF CANADA, *supra* note 252, at 2; UNITED STATES SENATE COMM. ON GOVERNMENTAL AFFAIRS, SUMMARY OF INTELLIGENCE REFORM AND TERRORISM PREVENTION ACT OF 2004 (Dec. 6, 2004), *available at* http://hsgac.senate.gov/files/ConferenceReportSummary.doc .

[260] UNITED STATES SENATE COMM. ON GOVERNMENTAL AFFAIRS, *supra* note 259.

among intelligence agencies to accomplish necessary purposes.[261] The DNI would be able to monitor compliance if agencies were required to submit privacy impact assessments or other privacy checks before receiving funding for data mining projects.

- Monitor heads of intelligence centers and agencies and has the authority to make personnel changes.

- Ensure quality of the intelligence being generated.[262] This will be an area where the DNI will constantly be required to balance the quality of the intelligence with potential privacy invasions. The DNI will have a Civil Liberties Protection Officer to counteract the urge to obtain the best quality, regardless of Fourth Amendment protections.[263]

### ii. PRIVACY AND CIVIL LIBERTIES BOARD

The drafters of the bill recognized that increased intelligence sharing would increase privacy concerns. Therefore, the IRTPA created a Privacy and Civil Liberties Board. The purpose of the Board is "to ensure that privacy and civil liberties concerns are appropriately considered in the implementation of laws, regulations, and executive branch policies related to efforts to protect the nation against terrorism."[264] For example, the Board will review the techniques used for the individual tracking systems used by border and immigration patrol, and the database used to track security clearances, both of which were authorized under IRTPA.[265]

This Board is responsible for reviewing policies and procedures in conjunction with the Privacy Office of the Department of Homeland Security and will be located in the Executive Office of the President.[266]

---

[261] *Id.*

[262] Cumming, *supra* note 251.

[263] UNITED STATES SENATE COMM. ON GOVERNMENTAL AFFAIRS, *supra* note 259.

[264] *Id.*

[265] *Id.*

[266] *Id.*; U.S. EMBASSY OF CANADA, *supra* note 252.

The Board will review a procedure put in place by the IRTPA that requires agencies to expend funds to ensure the security of information, as soon as a data technology project begins.[267] The Board is also charged with working with the President to create an Information Sharing Environment.[268] The President must create an Information Sharing Environment to "facilitate the sharing of terrorism information among all appropriate federal, state, local, tribal, and private sector entities, through the use of policy guidelines and technology."[269] The Board will be involved in this development to ensure that the Information Sharing Environment does not become an arching government surveillance tool or reach into the private space of U.S. citizens.

### iii. NATIONAL IDENTIFICATION STANDARDS

Another implemented recommendation of the 9/11 Commission was the institution of national identification standards.[270] Before the IRTPA, all identification, such as driver's licenses and personal identification cards, were at the sole discretion of the states.[271] The IRTPA does not take away the power to issue identification from states, but rather creates standards that states must meet, and gives the Secretary of Transportation the authority to make federal standards.[272] These standards may include what information must be on the identification cards and the documentation required to obtain the card.

Privacy advocates fear that these standards will place higher burdens on states to deal with privacy concerns. State systems may not have the level of security necessary to protect this data, which may increase the chances of identity theft and misuse of private information. These increased expectations may be overwhelming for state and local governments trying to balance the demand for efficient

---

[267] UNITED STATES SENATE COMM. ON GOVERNMENTAL AFFAIRS, *supra* note 259.

[268] *Id.*

[269] *Id.*

[270] Tatelman, *supra* note 250.

[271] *Id.* at 1.

[272] *Id.*

technology uses and expectations of privacy.[273] Privacy advocates also fear that these standards are the first step toward a national identification system, which could create a climate of excessive government surveillance.[274]

### iv. IMPACT AND CRITIQUE OF THE INTELLIGENCE REFORM AND TERRORISM PREVENTION ACT OF 2004

Since the passage of the IRTPA, many policy analysts have critiqued and discussed the impact of the IRTPA. Some praise the IRTPA for creating privacy and oversight protections for information sharing, while others argue that there are not sufficient safeguards in place to prevent an abuse of intelligence information.[275] Most appreciate that the Privacy and Civil Liberties Board will be helpful to ensure that privacy concerns are considered when drafting policy.[276] There are concerns that this Board will be too politically motivated and lack independence, because the positions are Presidentially appointed.[277] Privacy advocates are concerned that the IRTPA is "too lax on limits on domestic spying."[278] The Act lowers standards for the FBI to conduct secret surveillance.[279] This lowering of the standard seems contrary to the purposes of some of the privacy safeguards in the IRTPA, like the Privacy and Civil Liberties Board. Similarly,

---

[273] National Electronic Commerce Coordinating Council, *Enterprise Identity and Access Management: The Rights and Wrongs of Process, Privacy and Technology* 2 (Nov. 2003) *available at* http://www.ec3.org/Downloads/2003/EnterpriseIdentity.pdf.

[274] Susan Llewelyn Leach, *A Driver's License as National ID?*, CHRISTIAN SCIENCE MONITOR (Jan. 24, 2005) *available at* http://www.christiansciencemonitor.com/2005/0124/p11s02-ussc.html.

[275] Reclaim Democracy, *Evaluating the Intelligence Reform and Terrorism Prevention Act of 2004* (Dec. 27, 2004), *at* http://reclaimdemocracy.org/articles_2004/evaluation_2004_intelligence_reform.html; Letter from the American Civil Liberties Union to Congress (Dec. 6, 2004) *available at* http://www.aclu.org/SafeandFree/SafeandFree.cfm?ID=17154&c=206.

[276] Reclaim Democracy, *supra* note 275.

[277] Letter from the American Civil Liberties Union to Congress, *supra* note 275.

[278] *Id.*

[279] Reclaim Democracy, *supra* note 275.

there is a concern that the Information Sharing Environment could become a tool for data surveillance without boundaries.[280]

### C. OTHER GOVERNMENT DATA MINING INITIATIVES

A variety of other data mining initiatives have arisen recently through government agencies. The Department of Homeland Security was authorized in its creation under the Homeland Security Act to participate in data mining.[281] This authorization will certainly result in a number of projects to further the Department's anti-terrorism purposes. One such project is Incident Data Mart, which will assemble data from any federal, state, or local law enforcement log to "spot possible terrorist activities."[282] Privacy advocates will continue to monitor these activities to ensure Fourth Amendment protections are incorporated into their development.

Domestic law enforcement entities are increasingly using data mining. Recent focus has been put on developing and implementing the National Criminal Intelligence Sharing Plan. This is a plan designed under the presumption that "no single government agency – or government – can win the war on terrorism."[283] The purpose of the plan is to ensure that the law enforcement community is effectively sharing information.[284] This plan consists of a variety of initiatives that focus on different types and uses of information. Current initiatives include the Global Justice Information Sharing Initiative, Law Enforcement Information Sharing Initiative, Criminal Intelligence Coordinating Council, Criminal Justice Information Systems Division, as well as cooperation with the FBI.[285] This plan has been developed

---

[280] Letter from the American Civil Liberties Union to Congress, *supra* note 275.

[281] CENTER FOR DEMOCRACY AND TECHNOLOGY, *supra* note 240.

[282] Declan McCullagh, *Government Data-mining Lives On,* NEWS.COM (June 1, 2004), *available at* http://news.com.com/Government+data-mining+lives+on/2010-1028_3-5223088.html.

[283] Press Release, U.S. Dept. of Justice, Fact Sheet: National Criminal Intelligence Sharing Plan (May 14, 2004), *available at* http://www.fas.org/irp/news/2004/05/doj051404.html.

[284] *Id.*

[285] *Id.*

for the federal justice constituency, but can be adapted to fit other sizes and types of programs in need of an information sharing directive.[286]

The FBI has found use in data mining and employs it through several programs. It believes that the "sharing of intelligence [is] absolutely critical to prevention efforts."[287] One of the most recent programs is the National Intel Share Project.[288] The most current efforts are focused on gathering data from all FBI computers into a single, searchable data warehouse. Varying levels of sophistication in data mining are also being used by the Joint Terrorism Task Forces, the National Joint Terrorism Task Force, the Foreign Terrorist Tracking Task Force, and the Office of Law Enforcement Coordination.[289] Another criminal justice data mining project is the Financial Crimes Enforcement Network, which links all levels of law enforcement and financial communities.[290] As these programs are implemented and become more prevalent, they will need to be evaluated for privacy protection.

## III. STATE DEVELOPMENTS IN GOVERNMENT DATA MINING: MULTISTATE ANTI-TERRORISM INFORMATION EXCHANGE (MATRIX)

The federal government is not the only entity taking advantage of the benefits of data mining. State governments are implementing data mining techniques to further process improvements, law enforcement, and various other purposes served by analyzing large quantities of data. In the wake of September 11, 2001, states realized that they needed to play a role in fighting terrorism. The MATRIX program was designed specifically for this reason. As with the TIA, the MATRIX program has been subject to much criticism and decreased participation.

---

[286] GLOBAL INFRASTRUCTURE/STANDARDS WORKING GROUP, A FRAMEWORK FOR JUSTICE INFORMATION SHARING: SERVICE-ORIENTED ARCHITECTURE (SOA) 5 (Dec. 9, 2004), *available at* http://it.ojp.gov/process_links.jsp?link_id=4418.

[287] FEDERAL BUREAU OF INVESTIGATION, PARTNERSHIPS (2004), *available at* http://www.fbi.gov/terrorinfo/counterrorism/partnership.htm.

[288] *Id.*

[289] *Id.*; Robert A. Martin, *The Joint Terrorism Task Force: A Concept That Works*, Anti-Defamation League (2004), *available at* http://www.adl.org/learn/jttf/default.asp.

[290] SOURCEWATCH, *Financial Crimes Enforcement Network* (2005), *at* http://www.sourcewatch.org/wiki.phtml?title=Financial_Crimes_Enforcement_Network.

## A. MATRIX OVERVIEW

The MATRIX program was originally designed by Hank Asher of Seisint, Inc., for the state of Florida.[291] Seisint, a private database contractor, still operates the program which mines a large amount of information including: marriage and divorce data, criminal history records, names and addresses of family and neighbors, names and addresses of business associates, driver's license data, vehicle registration records, incarceration records, and digital photos.[292] The purpose of gathering this information is "to assist in investigations pertaining to threats to national security."[293] This centralized database pulls together local, state, and federal law enforcement data and intelligence, along with public databases from agencies like the department of motor vehicle and property records.[294] This information can then be searched according to a specific subject or through pattern analysis.

The program has a variety of uses for state law enforcement officials. Like other data mining programs, data is cross-linked to other data using information technology expertise, public and private data sets, and a supercomputer.[295] The software can run pattern-based inquiries to create watch lists and determine possible terrorism suspects.[296] State governments and Seisint claim this is not data mining, despite the scrutiny that it is more than just gathering data.[297]

---

[291] Interview by Robert O'Harrow with Hank Asher, Founder of Seisint (2004), *available at* http://www.noplacetohide.net/asher.html.

[292] SOURCEWATCH, *Multistate Anti-Terrorism Information Exchange Program* (2005), *at* http://www.sourcewatch.org/wiki.phtml?title=Multistate_Anti-Terrorism_Information_Exchange_Program.

[293] Letter from Marshal Home, Commissioner, Georgia Department of Motor Vehicle Safety to Jim Lientz, Chief Operating Officer, State of Georgia Office of the Governor 1 (Sept. 29, 2003), *available at* www.dmvs.ga.gov; *see also* http://www.matrix-at.org/.

[294] PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[295] Blumenthal, *supra* note 219.

[296] Anderson, *supra* note 203, at 6; PEOPLE FOR THE AMERICAN WAY, *supra* note 201.

[297] Press Release, American Civil Liberties Union, ACLU Unveils Disturbing New Revelations About MATRIX Surveillance Program (May 20, 2004), *available at* http://www.aclu.org/Privacy/Privacy.cfm?ID=15834&c=130.

They claim that it is "factual data analysis."[298] One specific function that MATRIX hallmarked was the High Terrorist Factor. This quotient was used "to measure the likelihood that individuals in the company's databases were terrorists."[299] After searching its records, based on this quotient, the database returned 120,000 people with high terrorist factors. After only approximately 40 of these led to arrests, the quotient was no longer used.[300]

The MATRIX program was designed for the use of state governments, rather than the federal government. Florida originated the program, with the backing of Governor Jeb Bush.[301] At least fifteen states had committed to the program by the beginning of 2003; however, various problems arising with the program, as discussed below, caused many states to cease their participation. Texas withdrew in May 2003 and California followed by withdrawing in June 2003.[302] As of July 25, 2003, thirteen states were committed to the MATRIX program.[303] Over the next year, eight more states would cease participation, leaving only five states participating as of May 2004. These states included Connecticut, Florida, Michigan, Ohio, and Pennsylvania.[304] In August 2004, the American Civil Liberties Union (ACLU) and former Governor William Milliken filed suit against the State of Michigan, contending that the state police's participation in the MATRIX program was a violation of the state's Interstate Law Enforcement Intelligence Organization's Act of

---

[298] *Id.* at 4; *see also* American Civil Liberties Union, *MATRIX: Myths and Reality* 2 ACLU PRIVACY & TECHNOLOGY (Feb. 10, 2004), *available at* http://www.aclu.org/Privacy/Privacy.cfm?ID=14894&c=130.

[299] American Civil Liberties Union, *supra* note 297, at 2.

[300] Tien, *supra* note 158, at 5.

[301] American Civil Liberties Union, *supra* note 297, at 1.

[302] Letter from Marshall Caskey, TX Chief Criminal Law Enforcement to Tim Moore, Chair of Project MATRIX, Florida Dept. of Law Enforcement (May 21, 2003); Letter from Jim T. Moore, Commissioner, Florida Dept. of Law Enforcement to Patrick Lunney, Director for CA Dept. of Justice (June 12, 2003).

[303] Letter from Andrew T. Mitchell, Director, Office for Domestic Preparedness to Emory Williams, Institute for Intergovernmental Research (July 25, 2003); Seisint, *MATRIX: First Responder Support* (Jan. 24, 2003), *available at* www.seisint.com.

[304] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148, at 5.

1980.[305] The ACLU and former governor contend that the law requires participation in such a program to be implemented only by the approval of the state legislature or with a citizen oversight body, in order to "prevent unsupervised and uncontrolled access to information about individuals."[306] Currently, Michigan is still a member of the MATRIX program.

## B. CRITIQUES OF THE MATRIX PROGRAM

Privacy advocates fear that MATRIX is a scaled down version of the TIA.[307] Opponents of the program believe that the same data mining techniques that would have been used in the TIA to provide Pentagon anti-terrorism experts with information can now be disseminated to any local law enforcement official.[308] The difference in levels of authority could also lead to differences in the level of information security. The ability for information abuse appears higher because of the lower level of expertise and training of the end users.

Others believe that MATRIX is a scaled down TIA program because of the source of its funding. Although states are required to pay for the program, it is primarily supported through $8 million from the Department of Homeland Security and $4 million from the Justice Department.[309] It is argued that the federal departments are funding this program in an effort to make it nationwide, thereby replacing the TIA.[310] This is supported by the fact that the Department of Homeland Security requires managerial oversight and control, and enters into a "Cooperative Agreement" rather than a grant as a prerequisite for

---

[305] Press Release, American Civil Liberties Union, ACLU of Michigan and Former Governor Charge State Police with Violating Data Collection Law Through Controversial MATRIX Database Program (Aug. 3, 2004), *available at* http://www.aclu.org/Privacy/Privacy.cfm?ID=16206&c=130.

[306] *Id.*

[307] Blumenthal, *supra* note 219.

[308] *Id.*

[309] Anderson, *supra* note 203, at 6; American Civil Liberties Union, *supra* note 297, at 4; Barry Steinhardt, *Privacy and The Matrix,* THE WASHINGTON POST (May 27, 2004) *available at* http://www.washingtonpost.com/wp-dyn/articles/A57156-2004May26.html.

[310] Blumenthal, *supra* note 219.

dispensing the $8 million.[311] This arrangement has come into question and the ACLU has asked Nuala O'Connor Kelly, the Department of Homeland Security Privacy Officer, to investigate the Department of Homeland Security's involvement in the state data mining MATRIX project.[312]

Not only do privacy advocates have problems with the MATRIX program, many states do as well. The recent decline in the number of participants indicates that the program is not the ideal centralized database system for the majority of states. First, state governments generally operate on tight budgets and the funding required from each state is not feasible.[313] Each state must pay an annual charge of $1.78 million to participate in the program, as well as commit sufficient personnel and training dollars to allow the program to run successfully.[314] Second, states already find it difficult to obtain information from its citizens for needed services. Many citizens refuse to provide digital signatures, fingerprints, or social security numbers when getting their driver's licenses.[315] If these individuals knew that the information would then become part of a large database for the government to learn more about them and investigate them for possible wrongdoing, many more would not give the requested information or would give false information. This would then lead to results based on inaccurate or missing data. Fourth, states are concerned about giving public data to a non-public entity, Seisint. States that have chosen not to participate believe that there are "legal, ethical, and financial considerations in providing non-public data sets at [the state's] expense to a private company to sell back to [the state]."[316] The use of only one company essentially creates a

---

[311] Press Release, American Civil Liberties Union, ACLU Unveils Disturbing New Revelations About MATRIX Surveillance Program (May 20, 2004), *available at* http://www.aclu.org/Privacy/Privacy.cfm?ID=15834&c=130; *see also* American Civil Liberties Union, *supra* note 297, at 4; Steinhardt, *supra* note 309.

[312] American Civil Liberties Union, *supra* note 311.

[313] Letter from Marshall Caskey, Texas Chief Criminal Law Enforcement to Tim Moore, Chair of Project MATRIX, Florida Dept. of Law Enforcement, *supra* note 302, at 1.

[314] Letter from Marshal Home, Commisioner, Georgia Department of Motor Vehicle Safety to Jim Lientz, Chief Operating Officer, State of Georgia Office of the Governor, *supra* 293, at 2.

[315] *Id.*

[316] Letter from Marshall Caskey, TX Chief Criminal Law Enforcement to Tim Moore, Chair of Project MATRIX, Florida Dept. of Law Enforcement, *supra* note 302, at 2.

monopoly for Seisint, because it controls the database and the data. The system is likely not compatible with other systems and would be difficult, if not impossible, to transfer. Therefore, a state using the MATRIX program would have to continue to use Seisint, thereby creating a monopoly in violation of public policy.[317]

Most of the concerns with data mining generally apply to the MATRIX program, like lack of safeguards, inaccurate data collection, and no procedures for ensuring accuracy.[318] The Fourth Amendment issues continue to arise with arguments that the systems should only be used after probable cause is generated, in order to avoid treating all Americans as suspects to be eliminated.[319] There is even greater concern at the state level that MATRIX is being promoted to the public as an anti-terrorism tool, but will actually be used as a general law enforcement tool.[320] This perception may be greater at the state level because citizens are more likely to view terrorism as a national problem and do not see the states playing a large role in terrorism prevention. A more specific concern is the possibility of the program discriminating against certain ethnic groups based on the patterns that are searched. For example, it is likely that Middle Eastern men will be targeted based on the search methods of the MATRIX. This was especially hazardous when the High Terrorism Factor index was used.[321]

## IV. CONTROVERSIES IN GOVERNMENT DATA MINING

It is agreed that technology is useful, but needs to be designed in a way to protect privacy.[322] There must be methods to limit or restrict data mining to protect privacy, while maintaining or enhancing the technology's usefulness. Ways to achieve this balance have been the topic of much academic and professional research. While this balance is being sought, there will be legal disputes about its constitutionality and its misuse.

---

[317] Steinhardt, *supra* note 309.

[318] American Civil Liberties Union, *supra* note 305.

[319] Associated Press, *supra* note 217; Steinhardt, *supra* note 309.

[320] American Civil Liberties Union, *supra* note 298.

[321] American Civil Liberties Union, *supra* note 297, at 2.

[322] Dempsey & Rosenzweig, *supra* note 179, at 1; Rosenzweig, *supra* note 204.

### A. OBTAINING BALANCE BETWEEN PRIVACY PROTECTIONS AND USEFUL DATA MINING

A variety of suggestions have been made to improve data mining technology. The recommendations vary in degree of severity. One of the more intense suggestions is to use data mining only when there is probable cause and the investigation is so imminent that only "aggressive preventative strategies" are required.[323] However, more moderate proposals have been put forth. These suggestions deal with changes in technology, increasing transparency, and requiring greater third party oversight.

First, more secure technology has been strongly suggested. Three technology features that could promote privacy are anonymization of data, required authorization for access, and built-in audit logs. By anonymizing the data, the information could be accessed, but a specific name or identification would not be placed with it until necessary.[324] The benefit of this would be that personally identifiable information does not flow between sources. The information is not labeled or attached to a particular name. This would reduce the possibility of the identity theft and information misuse problems. The technology used to anonymize data is referred to as "one-way hashing," which scrambles the information so it can not be easily connected.[325] Access restriction can be used to prevent unwanted users from obtaining information.[326] Database access could be even further restricted so that users would be required to enter the purpose of their search, and would only receive related data and results while all other information would be restricted.[327] Audit logs are often used to track user access with many different technologies. They could be used in data mining systems to ensure that users were not accessing inappropriate information or otherwise abusing the system. This would require a neutral party to review the logs to ensure that the searches were not outside the bounds of appropriateness.

---

[323] Taipale, *supra* note 152, at 5.

[324] Dempsey & Rosenzweig, *supra* note 179, at 8; Associated Press, *supra* note 217.

[325] Dempsey & Rosenzweig, *supra* note 179, at 8.

[326] Associated Press, *supra* note 217; Dempsey & Rosenzweig, *supra* note 179, at 32.

[327] Dempsey & Rosenzweig, *supra* note 179, at 13.

Second, increased transparency of how the process works and what information is being mined would increase the privacy that some feel is violated by using these programs. Currently, the general public does not have information on the processes used by the government to obtain information on citizens; therefore, there is little debate and input from those outside of the system creators and a select number of policy analysts.[328] Without this input, the designers do not have all the necessary information and cannot make the most informed decisions on how to design the data mining program to protect the citizens' privacy interests. It has been suggested that the Fair Information Practices should apply to data mining.[329] There should be notification that the information is being used, a limit on how long the information can be retained, and the quality of the data must be examined.[330] These practices, specifically the notification requirement, would increase the transparency of data mining projects to the public. The implementation of these suggestions may be highly dependent on cost and logistical practicability.

Third, neutral oversight of data mining projects could be helpful in balancing valuable technologies in data mining against privacy protection. Congressional review could provide a layer of accountability before a data mining project is funded or implemented by a federal agency.[331] Large data mining users, such as the Department of Defense, may wish to apply administrative procedures that must be followed when it uses data mining.[332] The Privacy and Civil Liberties Board may wish to determine a model set of procedures for agencies and departments to follow. Another form of review would be to require judicial review of data mining programs. Courts would determine whether the process was in accordance with due process, and would be especially useful when the data mining involves personally identifiable information.[333]

---

[328] Anderson, *supra* note 203, at 16.

[329] CENTER FOR DEMOCRACY & TECHNOLOGY, PRIVACY'S GAP: THE LARGELY NON-EXISTENT LEGAL FRAMEWORK FOR GOVERNMENT MINING OF COMMERCIAL DATA 13 (May 28, 2003), *available at* http://www.cdt.org/security/usapatriot/030528cdt.pdf.

[330] Dempsey, *supra* note 176, at 9.

[331] Seifert, *supra* note 160, at 13; Taipale, *supra* note 152, at 5.

[332] Anderson, *supra* note 203, at 14.

[333] Pear, *supra* note 165; Taipale, *supra* note 152, at 5.

### B. FUTURE LAWSUITS REGARDING GOVERNMENT DATA MINING

Few lawsuits have arisen despite the controversy of data mining. However, as data mining becomes more common, legal disputes will undoubtedly surface. Two types of disputes may dominate this field. First, there will likely be requests to restrict admissibility of evidence obtained solely through data mining.[334] These court requests will be based on the theory that the evidence was obtained without probable cause and constituted an unreasonable search, thereby violating the Fourth Amendment. Second, corrective lawsuits may be brought seeking damages caused by misidentification. For example, if an individual is placed on a no fly list or a watch list and this causes damages, like not obtaining a job or losing a business contract, the person will likely be able to seek restitution for the injury. These potential legal challenges should prompt data mining initiatives to continue to work on minimizing privacy concerns and the use of accurate data.

## V. GOVERNMENT DATA MINING FROM PRIVATE INSTITUTIONS

The government collects a vast amount of data on its own. However, it also recognizes that private institutions in the United States gather data from individuals during daily business activities that could be useful for government purposes. Therefore, the government obtains data and data mining services from commercial providers. This extra data can give the government more information about one individual or more parameters with which to do pattern searches. There are less regulations on data collection by private institutions than by the government, so there are concerns that the government uses private institutions to avoid these restrictions.

### A. USE OF INFORMATION FROM PRIVATE INSTITUTIONS

The government can easily find commercial data aggregators from whom to purchase data or data mining and all levels of government are willing to do so.[335] According to the 2004 GAO report, fifty-four government data mining projects use private sector data. Of these projects, thirty-six involve the government obtaining personal data,

---

[334] Anderson, *supra* note 203.

[335] U.S. GENERAL ACCOUNTING OFFICE, *supra* note 148.

such as social security numbers and driver's license numbers, from commercial information.[336] The Immigration and Naturalization Service "queries private sector databases 20,000 times a month."[337] Not only is the government using these databases, but the usage is increasing. The FBI's use of commercial database information grew 9,600 percent between 1992 and 2003.[338]

Although purchasing data from commercial sources is common, government agencies and commercial databases may have incentive not to share information. If the government is working on an investigation of top priority, it may wish to avoid commercial sources in order to preserve the secrecy of its operation. Through the request of certain information, the provider may be able to deduce who the individuals of interest are and who is being investigated. This may enable commercial aggregators or their employees to jeopardize investigations by tipping off targets.[339] On the contrary, commercial databases may not wish to provide their information to the government. Businesses worry that if customers know that their information may be provided to the government, they will give false information.[340] This will ultimately reduce the value of the results for both the commercial user and the government.

### B. INAPPLICABILITY OF THE FOURTH AMENDMENT AND THE PRIVACY ACT OF 1974

The Fourth Amendment protects the privacy of individuals against government intrusion, but not against the private sector. Therefore, a conflict arises when the government obtains information that was gathered by commercial aggregators, as to whether there is a Fourth Amendment violation. Private companies are not accountable to the same extent that the government is, so there are concerns that the government can use this process as a backdoor method of obtaining

---

[336] *Id.* at 10.

[337] Declan McCullagh, *JetBlue Privacy – Under Federal Wings?*, NEWS.COM (Sept. 23, 2003), *at* http://news.com.com/JetBlue+privacy--under+federal+wings/2010-1029_3-5080339.html?taag=nl.

[338] Dempsey, *supra* note 176, at 10.

[339] *Id.* at 7.

[340] *Id.* at 7.

information that it could not collect itself.[341] Consumers do not have an expectation of privacy regarding the information that they divulge to businesses.[342] The Supreme Court has stated that the individual has no protection under the Constitution because this information was knowingly exposed.[343] However, some argue that this decision was given before the rapid expansion in technological capabilities and the ease and efficiency of sharing data could not have been foreseen by the Court at that time.[344]

The Privacy Act of 1974 provides little restraint on the data mining from commercial aggregators.[345] The government can obtain data mining services from the private sector without invoking the Privacy Act because the information is not actually part of the government's database.[346] Many see this as a loophole of the Privacy Act that allows the government to have full use of information that it cannot legally obtain on its own.[347] They believe that at the time of the Privacy Act, Congress could not anticipate the growth in information technology, which is the reason for not placing restrictions on obtaining information from private sources.[348]

A recent investigation into the role of the Privacy Act of 1974 in private sector and government information sharing resulted from a data transfer from the airline, JetBlue. JetBlue gave personal records of five million customers in early 2002 to Torch Concepts, a private contractor that provides data mining services to the Department of Defense.[349] This deal was negotiated between JetBlue and the

---

[341] Peter Jennings, *supra* note 157.

[342] CENTER FOR DEMOCRACY & TECHNOLOGY, *supra* note 329, at 2; Anderson, *supra* note 203, at 14.

[343] CENTER FOR DEMOCRACY & TECHNOLOGY, *supra* note 329, at 2.

[344] Dempsey, *supra* note 176, at 10.

[345] CENTER FOR DEMOCRACY & TECHNOLOGY, *supra* note 329, at 1; Anderson, *supra* note 203, at 14.

[346] CENTER FOR DEMOCRACY & TECHNOLOGY, *supra* note 329, at 3.

[347] Anderson, *supra* note 203, at 14.

[348] McCullagh, *supra* note 337.

[349] Blumenthal, *supra* note 219; *Id.*

Transportation Security Administration.[350] After a thorough investigation, the Department of Homeland Security Privacy Office found that there was not a violation of the Privacy Act by the Transportation Security Administration. However, it did specify that obtaining the data by negotiating directly with the data owner was not within the spirit of the Privacy Act.[351]

In order to protect citizens from a backdoor invasion of privacy, several changes have been recommended. A stringent change would be to adapt the Privacy Act and the Fair Information Practices to apply to commercial databases.[352] Many feel that because these rules are focused on upholding the Fourth Amendment right against government intrusion, that they, as well as other restrictions, should not be applied to the private sector. Instead, restrictions should be placed on what information the government can obtain and how it can obtain it.[353] For example, government agencies could be required to disclose which private sector databases they have purchased and will be using.[354] Another suggestion is that government agencies undergo Privacy Impact Assessments before engaging in any data mining.[355] These assessments would require a close look at the source of the data and the potential invasions of privacy rights that may occur if the data mining occurs.

## VI. THE FUTURE OF GOVERNMENT DATA MINING AND CONCLUSIONS

It is undeniable that most data mining techniques are useful processes that must be considered as viable and efficient methods for investigation and problem solving. It is also undeniable that most data mining invokes a certain level of privacy concerns. Finding the delicate balance between privacy and usefulness will be a goal for government agencies and privacy advocates to work out. As

---

[350] DEP'T OF HOMELAND SECURITY PRIVACY OFFICE, REPORT TO THE PUBLIC ON EVENTS SURROUNDING JETBLUE DATA TRANSFER 9 (Feb. 20, 2004), *at* http://www.epic.org/privacy/airtravel/jetblue/dhs_report.pdf.

[351] *Id.* at 9.

[352] Dempsey, *supra* note 176, at 14.

[353] McCullagh, *supra* note 337.

[354] *Id.*

[355] Dempsey, *supra* note 176, at 8.

technology continues to advance, there will be additional likelihood that the balance will continually be in flux. The Privacy and Civil Liberties Board, Director of National Intelligence, and Department of Homeland Security Privacy Officer, among others, will be responsible for putting procedures and safeguards in place that will allow the government to stay abreast of these changes.

It is uncertain as to whether changes will be made in the TIA to make it acceptable to receive additional Congressional funding. Most likely, the technologies being developed by the TIA will continue to be disseminated to other government agencies and promoted for various purposes, some of which will be similar to the anti-terrorism purposes of the TIA. The Director of National Intelligence must be appointed, according to the Intelligence Reform and Terrorism Prevention Act of 2004. This director will likely evaluate the current intelligence communities and determine how and to what extent data mining will be used. These data mining efforts will be more conscious of privacy protections due to the controversy over the TIA and the implementation of the Privacy and Civil Liberties Board. It is likely that privacy advocates will closely watch the Information Sharing Environment that is being created under this bill to ensure that it does not pose privacy threats similar to those found in the TIA.

On the state level, the MATRIX program will likely continue to decline as states realize not only the privacy concerns, but the budget strain that the program imposes. States may be more likely to rely on federal government initiatives to fight terrorism and limit data mining to internal uses like service and process improvements. However, state governments may continue to use the services of outside contractors, like Seisint, to obtain data for the functions that will require data mining. Both state and local governments will continue to use private contractors with private databases to obtain larger amounts of data and save money on system development and data collection. As this trend increases, Congress will likely propose amendments to the Privacy Act of 1974 that will impose accountability obligations and regulations on government entities that obtain information from private data mining sources.

The trend toward the increasing use of data mining will likely continue, as organizations, both commercial and public, realize the capabilities of such programs and the opportunities it provides in many areas, not simply in anti-terrorism and law enforcement. However, these progressions will likely be faced with increased governmental control in the form of regulations, congressional action, and potential judicial review. Governmental data mining will continue to be a source of debate and development as technological advances must find their fit within the Fourth Amendment privacy protections.

## CONCLUSION

New technologies allow government agencies to collect more personal information about citizens. Every time a new technology comes into place, new kinds of privacy concerns are raised. Government agencies often need to promote and adopt new technologies to ensure citizen safety on airplanes and elsewhere. They also need to encourage citizen to use the electronic tools. Congress provides extra measures to protect individual privacy when the government agencies start new electronic programs, even though they do not always address all concerns raised by the public.

The new laws mandating PIA and GAO's authorization before government agencies implement new electronic programs ensure that agencies consider their privacy implications, and inform the public of their privacy rights when they collect information about individuals. On government websites, all federal agencies are required to notify the public about their privacy policies, and many states follow the same practice. Even with these laws, however, the public is still concerned about its privacy when the government collects individual information from airline passenger records. Public interest organizations, like the ACLU and the National Business Travel Association, ask the agencies to modify some of the controversial agency practices, such as the use of commercial databases to identify persons with a high risk of conducting terrorism and lack of proper redress systems under the CAPPS II. These concerns eventually led the TSA to abandon the program in 2004. Instead, Secure Flight was created to answer most of these concerns. However, this program also was the subject of privacy concerns from the public. It is not clear how the TSA will ensure that the new program is properly implemented step-by-step in accordance with the PIAs and other laws. In order to obtain public approval, the TSA and other government agencies need to balance better security with policies and procedures that protect privacy.

Data mining has established itself as a valuable product for a variety of commercial and government purposes. Therefore, state and local governments will need to determine how to gather information, design data mining programs, and implement these programs in a fashion that accounts for the concerns of individual citizens. Programs like the TIA will continue to be developed but can benefit by considering some of the technological, transparency, and oversight suggestions that can protect the information mined and the knowledge gathered from data mining systems. The privacy oversight officials and boards put into place by the Intelligence Reform and Terrorism

Prevention Act of 2004 will be responsible for ensuring that either these safeguards or more effective techniques are used. The MATRIX is likely to continue to be scrutinized for its federal involvement, and may lose additional members based on the large expense the program to states. Amendments to the Privacy Act of 1974 have the potential to influence the way that the government accesses data from private contractors. The data mining versus privacy debate may only be starting, and additional concerns will continue to appear and be addressed with each new data mining program and new technologies developed.

As analyzed in this article, new government programs can intrude on the privacy of citizens. Many citizens and public interest organizations see the need of extra security to make the country safe in the aftermath of the September 11, 2001 attacks. However, privacy is one of the most important and fundamental rights of citizens and is protected under the Fourth Amendment of the Constitution. Therefore, the government must ensure that it balances better security with privacy protection. In the future, new technologies would allow government agencies to access more diverse information about citizens. Each time an agency adopts a new technology or new method of collecting information, it needs to consider if privacy will be protected, and ensure that there is an option either to abandon the new measure or provide an additional law or other measures to secure privacy.